

Сканирование и форматирование: [Янко Слава](#) (Библиотека [Fort/Da](#)) || [slavaaa@yandex.ru](mailto:slavaaa@yandex.ru) ||  
[yanko\\_slava@yahoo.com](mailto:yanko_slava@yahoo.com) || <http://yanko.lib.ru> || Исq# 75088656 || Библиотека:  
<http://yanko.lib.ru/gum.html> || Номера страниц - внизу  
update 09.05.06

---

РОССИЙСКАЯ АКАДЕМИЯ НАУК ИНСТИТУТ ФИЛОСОФИИ

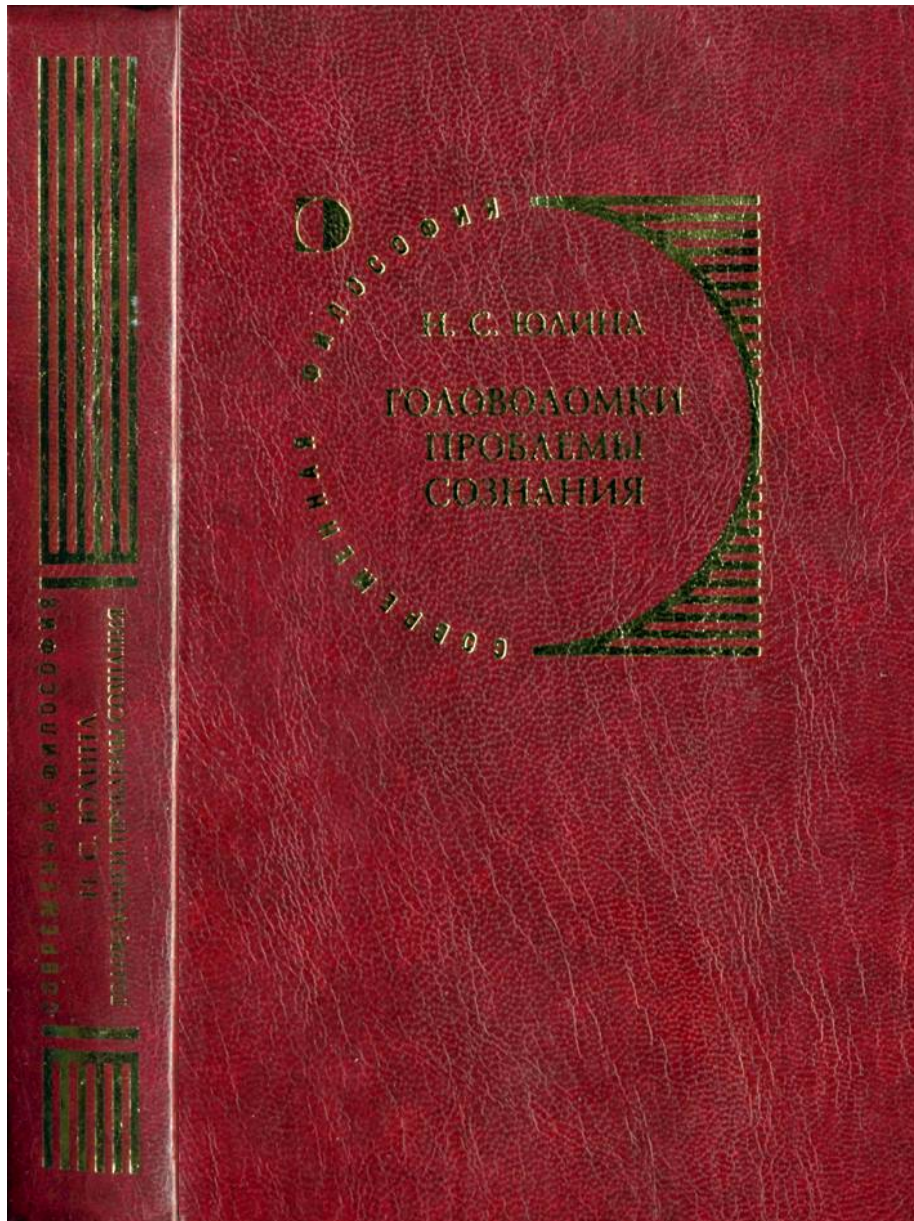
**Н. С. ЮЛИНА**

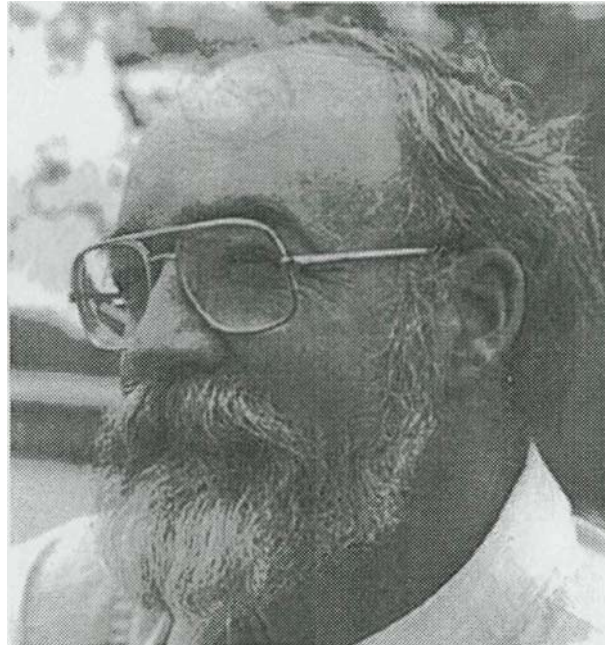
## **ГОЛОВОЛОМКИ ПРОБЛЕМЫ СОЗНАНИЯ**

концепция Дэниела Деннета

*с цв. тиснением  
от автора  
Н. С. Юлиной  
25.05.2005*

КАНОН +  
Москва 2004





УДК 14 ББК 87.9 Ю 17

Утверждено к печати Ученым советом Института философии РАН

Настоящее издание осуществлено при финансовой поддержке Российского гуманитарного научного фонда (РГНФ), проект № 03-03-00312д.

Рецензенты: д. ф. н. *Дубровский Д. И.*, д. ф. н. *Зотов Л. Ф.*

**Юлина Н. С.**

**Ю 17 Головоломки проблемы сознания:** концепция Дэниела Деннета. — М.: Канон+, 2004. - 544 с. — (Современная философия).

ISBN 5-88373-224-0.

В книге рассказывается о концепции сознания и самости Дэниела К. Деннета, одного из ярких представителей современной философской мысли в США. Концепция интересна своей попыткой дать последовательно материалистическое (функционалистское) объяснение головоломкам проблемы сознания с применением дарвиновской теории, а также подходов, используемых в когнитивных науках и исследованиях по искусственному интеллекту. Автор рассматривает идеи Деннета в контексте дискуссий англоязычной философии и сложившихся в ней альтернатив в объяснении сознания. Во второй части книги сопоставляется позиция Деннета с позициями его оппонентов (Т. Нагель, Дж. Сёрль, Р. Рорти, К. Поппер и др.). Содержание книги может представлять интерес для философов, психологов, специалистов по искусственному интеллекту и всех тех, кого интересует состояние современной мысли в США.

This book is about conception of consciousness and the self by Daniel Dennett a contemporary well-known and outstanding philosopher in the USA. His conception is an attempt to give a consistently materialistic (functionalistic) explanation of the puzzles of mind, applying Darwin's ideas and tools of cognitive sciences and artificial intellect. Dennet's ideas are considered in the context of the discussions in current Anglo-American philosophy and alternativ explanations of consciousness. The second part of the book gives a comparison of his position and those of T. Nagel, J. Searle. R. Rorty, K. Popper and others.

УДК 14

ББК 87.9

ISBN 5-88373-224-0

© Юлина Н. С. © Издательство «Канон +» ОИ «Реабилитация», 2004

## Электронное оглавление

Электронное оглавление .....	5
<b>Часть I. ГОЛОВОЛОМКИ ПРОБЛЕМЫ СОЗНАНИЯ .....</b>	<b>8</b>
<b>I. Полвека атак на «БАСТИОН СОЗНАНИЯ» .....</b>	<b>8</b>
1. О Дэниеле Деннете .....	8
2. Презентация философии Деннета: трудности ее исследования .....	12
3. Тайна феномена сознания .....	16
4. История атак на дуализм .....	20
<b>II. Деконструкция опыта и сознания .....</b>	<b>29</b>
1. Иллюзии «Картезианского Театра» .....	29
2. «Картезианский материализм» .....	32
3. Миражи «феноменологического сада» .....	34
4. «Дисквалификация квалиа» .....	38
5. Иметь опыт — значит иметь суждение об опыте .....	44
6. Почему нельзя создать чувствующий боль компьютер? .....	46
7. Животные и младенцы, не имеющие языка, не имеют сознания .....	50
Гетерофеноменология или как исследовать сознание эмпирически .....	54
<b>III. Сознание в рамках «Модели Множественных Набросков» .....</b>	<b>58</b>
1. От теории тождества к функционализму «Машины Тьюринга» .....	58
2. «Модель Множественных Набросков»: децентрализация сознания .....	66
3. Аномалии восприятия времени: «сталинская» и «оруэлловская» модели его редактирования .....	69
Поиск значений: обратная связь невербального и вербального .....	71
<b>IV. «Машина Тьюринга», «Машина Дарвина» и виртуальная «Машина Сознания» .....</b>	<b>76</b>
1. Компьютерные метафоры и философия .....	76
Виртуальная «Машина Сознания»: ее «размещение» в органике мозга .....	79
3. Эволюционизм, язык и роль меме-вирусов в производстве сознания .....	83
<b>V. Самость как «центр нарративной гравитации» или почему возможны самостные компьютеры .....</b>	<b>89</b>
1. Биологическая самость и нарративная самость .....	89
2. Почему возможно помыслить самостный компьютер .....	93
3. Свобода воли и детерминизм .....	94
4. Эволюционные алгоритмы и творчество .....	96
5. Где находится мое Я? .....	99
Телепортация и бессмертие .....	99
<b>VI. Интенциональность как установка рациональности .....</b>	<b>101</b>
1. Бихевиоризм и интенциональность .....	101
2. Физическая установка, установка дизайнера и интенциональная установка .....	104
3. Интенциональность, эволюционный аргумент и «свободно плавающая рациональность» .....	108
4. Уровни интенциональности и интеллектуальности .....	114
5. Первичная и производная интенциональность. Мать-Природа как первичный дизайнер .....	116
6. Джерри Фодор о телеологизме эволюционного аргумента Деннета .....	118
<b>VII. Об условиях личностного .....</b>	<b>123</b>
1. Необходимые и достаточные условия понятия «личностное» .....	123
2. «Личностное» в свете интенциональной установки .....	126
3. Нормативно личностное и метафизически личностное .....	128
<b>VIII. О ПОНЯТИИ «ОТВЕТСТВЕННОСТЬ» В СВЕТЕ МЕХАНИЦИСТСКОГО ОБЪЯСНЕНИЯ ЧЕЛОВЕКА .....</b>	<b>130</b>
1. Можно ли включить целевое объяснение человека в физикалистскую картину мира? .....	130
2. Совместимость механицизма и ответственности в свете теории установок .....	132
3. Действия, манипуляция и ответственность: каузальные и рациональные факторы .....	136
<b>Часть II. ДЕННЕТ И КРИТИКА .....</b>	<b>142</b>
<b>I. Споры вокруг философии Деннета .....</b>	<b>142</b>
<b>II. НАГЕЛЬ VERSUS ДЕННЕТ: «ОБЪЕКТИВНАЯ ФЕНОМЕНОЛОГИЯ» ПРОТИВ ПАНОБЪЕКТИВИЗМА .....</b>	<b>147</b>
<b>III. Сёрль versus Деннет: «Аргумент Китайской Комнаты» .....</b>	<b>151</b>
1. Материализм и «онтология субъективного» .....	151



2. Когнитивизм компьютера и человеческое понимание. Синтаксис и семантика.....	154
<b>IV. Деннет versus Рорти: МЕТАФОРИЗМ, РЕАЛИЗМ И «БАЗИСНЫЙ ИНСТИНКТ» .....</b>	<b>160</b>
Рорти о Деннете .....	162
1. «Коперниканский переворот» в философии сознания: от эссенциализма к прагматизму .....	162
2. Деннет «сидит на заборе» между реализмом и ирреализмом.....	165
Деннет о Рорти.....	168
3. О «привилегированных» репрезентациях старого когнитивизма и «непривилегированных» репрезентациях нового когнитивизма .....	168
4. Метафоризм, базисный инстинкт и «вегетарианская» концепция истины .....	171
5. Регулятивный идеал философии: объективность или солидарность? .....	173
6. О вирусе постмодернизма и «западной» науке .....	175
<b>V. К. Поппер и Д. Деннет: архитектура сознания согласно «Открытой» и «Закрытой» Вселенной.....</b>	<b>179</b>
<b>Заключение .....</b>	<b>186</b>
Заметки историка философии .....	186
<b>Приложение 1 Тексты Д.К. Деннета (Перевод с английского Н.С. Юлиной) .....</b>	<b>196</b>
1. По следу Дарвина. Где есмь Я? .....	196
2. Почему каждый из нас является новеллистом .....	210
3. Условия личностного.....	219
4. Почему нам важно понимать это правильно. Постмодернизм и истина .....	234
Истина против преходящих фантазий .....	236
Базисный инстинкт .....	238
Как доминирует наука .....	241
<b>Приложение 2. Библиография .....</b>	<b>243</b>
Книги Д.К. Деннета .....	243
Тексты Деннета на русском языке.....	243
Список статей, рецензий, интервью Деннета* (Работы с 1987 по 2003 гг. доступны в Интернете. ....)	243
2003 .....	243
2002 .....	244
2001 .....	244
2000 .....	244
1999 .....	244
1998 .....	245
1997 .....	245
1996 .....	245
1995 .....	246
1994 .....	246
1993 .....	247
1992 .....	247
1991 .....	247
1990 .....	248
1989 .....	248
1988 .....	248
1987 .....	248
1986 .....	249
1985 .....	249
1984 .....	249
1983 .....	250
1982 .....	250
1981 .....	250
1980 .....	250
1979 .....	250
1978 .....	251
1977 .....	251
1976 .....	251
1975 .....	251
1974 .....	251
1973 .....	251

1972 .....	251
1971 .....	251
1968 .....	251
<b>Литература о философии Деннета на русском языке.....</b>	<b>252</b>
<b>Литература о философии Деннета на английском языке.....</b>	<b>252</b>
Отдельные издания .....	252
Статьи .....	252
<b>Общая литература по проблеме сознания на английском языке .....</b>	<b>254</b>
<b>Общая литература о сознании на русском языке .....</b>	<b>257</b>
<b>Указатель имен* .....</b>	<b>260</b>
<b>Содержание .....</b>	<b>264</b>
<b>Contents .....</b>	<b>267</b>
<b>Приложения www = SL .....</b>	<b>269</b>
<b>Тексты Деннета из www .....</b>	<b>269</b>
In Darwin's Wake, Where am I? - December 29, 2000 .....	269
Postmodernism and Truth. FINAL DRAFT August 13, 1998 .....	279
Intentionality (1987) .....	286

*Светлой памяти моего отца  
Юлина Степана Семеновича  
посвящаю эту книгу*

## **Часть I. ГОЛОВОЛОМКИ ПРОБЛЕМЫ СОЗНАНИЯ**

### **I. Полвека атак на «БАСТИОН СОЗНАНИЯ»**

Современную философию сознания невозможно обсуждать без незримого присутствия духа  
Декарта.

*Якко Хинтикка*

Я склоняюсь к мысли, что в последнее время философия сознания является одной из немногих областей, в которой мы — профессора философии — делаем что-то стоящее.

*Ричард Рорти*

#### **1. О Дэниеле Деннете**

Учась в колледже, вспоминает Деннет, он прочитал «Meditations» Декарта, и в его душу на всю жизнь запал вопрос: «Как это вообще может быть, что мои мысли и чувства сосуществуют в том же самом мире с нервными клетками и молекулами, из которых состоит мой мозг?»<sup>1</sup> На протяжении всей творчес-

<sup>1</sup> Dennett D. *Consciousness Explained*. Boston, 1991. P. XI.

5

кой карьеры он пытается ответить на него, обращаясь к вееру разнообразных подпроблем. Можно ли говорить об особой реальности сознания и ощущений? Сможет ли нейроученый когда-либо прочитать мысли в моем мозгу? Что такое самость, свобода воли, личностное — иллюзии или реальность? Может ли робот в подлинном смысле быть наделен сознанием и самостью? Почему компьютер не чувствует боль? Совместимо ли целеполагающее и механистическое объяснение личности? Одна из работ Деннета называется «Brainstorms»<sup>1</sup>, что, с учетом содержания книги и его склонности к многозначным образам, на русский язык можно перевести и как «головоломки», и как «потрясения мозгов», и как «мозговой штурм», и как «появление плодотворных идей». Собственно говоря, все эти смыслы приложимы к его творчеству. Практически все оно посвящено разгадыванию головоломок феномена сознания и выдвиганию неординарных гипотез, не боясь, что у кого-то это вызовет «потрясение мозгов».

Дэниел Клемент Деннет родился в 1942 году в Ливане (его отец работал здесь в представительстве США), учился в Гарвардском и Оксфордском университетах. В Гарварде его учителем был Уилард Куайн, ставший для него другом и интеллектуальным героем, на которого он равнялся. В Оксфорде ментором при подготовке докторской диссертации был Джильберт Райл, стимулировавший его искать неординарные пути к объяснению сознания. В настоящее время Деннет является членом Американской академии, профессором университета Тафтса, директором «Центра когнитивных наук», соредактором влиятельного журнала «Behavioral and Brain Sciences», почетным профессором искусства и наук, читает лекции во многих университетах мира.<sup>2</sup> Русскоязычному читателю это

<sup>1</sup> Dennett D. *Brainstorms*. *Philosophical Essays on Mind and Psychology*. ch. 14. Montgomery, 1978. Sec. Ed. 1986.

<sup>2</sup> В середине 80-х годов в составе группы философов из США Деннет принял участие в проводившемся в Институте философии АН СССР семинаре на тему «Духовное и телесное».

6

имя малоизвестно<sup>1</sup>. Между тем он автор девяти книг, являющихся на сегодняшний день, пожалуй, наиболее фундаментальными и обсуждаемыми в философии сознания: «Содержание и сознание», «Головоломки. Философские очерки о сознании и психологии», «Пространство для свободы. Варианты свободы воли, стоящие того, чтобы их желать», «Интенциональная установка», «Сознание объясненное», «Опасная идея Дарвина. Эволюция и смысл жизни», «Виды психики»,



«Дети мозга», «Свобода развивается»<sup>2</sup>. Высказанные в них идеи активно дискутируются в философской среде, а также психологами, биологами, специалистами по когнитивным наукам и искусственному интеллекту. Две его книги («Сознание объясненное» и «Опасная идея Дарвина») попали в разряд бестселлеров 90-х годов, что большая редкость для философской литературы. (Последняя книга вызвала широкие дебаты и накалила остроту ведущихся в англоязычной литературе «дарвиновских войн».) Широкий отклик на идеи Деннета в немалой степени объясняется стилистикой его работ:

<sup>1</sup> На русском языке имеются следующие публикации: Деннет Д. Онтологическая проблема сознания // Аналитическая философия: становление и развитие. М., 1998. С. 361—375; Деннет Д. Условия личностного // История философии. М., ИФРАН, 2000. С. 199—223; Деннет Д. Постмодернизм и истина. Почему нам важно понимать это правильно // Вопросы философии, 2001. № 8. С. 93—100; Д. Деннет о проблеме ответственности в свете механицистского объяснения человека // История философии. ИФРАН. М., 2002. С. 58—77; Деннет Д. Виды психики. На пути к пониманию сознания. М., 2004.

<sup>2</sup> Content and Consciousness. L., Boston, 1969; Brainstorms. Philosophical Essays on Mind and Psychology. Montgomery, 1978; Elbow Room: the Varieties of Free Will Worth Wanting. Oxford, 1984; The Intentional Stance. L., 1987; Consciousness Explained. Boston, 1991; Darwin's Dangerous Idea. Evolution and the Meaning of Life. L., 1995; Kinds of Mind. Toward an Understanding of Consciousness. N.Y., 1996; Brainchildren. Essays on Designing Minds. Cambridge MA, 1998; Freedom Evolves, 2003.

7

неистощимая фантазия, образный разговорный язык, остроумные примеры, необычные аналогии, а также поразительная эрудиция в самых различных областях знания обеспечивают ему внимание научной публицистики и широкой читательской аудитории. Сегодня каждая его работа по выходе тут же переводится на основные языки мира, включая восточные.

По своим установкам Деннет одновременно и радикал, и консерватор. Он радикал в том, что настаивает на необходимости сквозного переосмысления образа «Вселенной Сознания» и деконструкции привычных представлений о ментальности, о наличии у нас внутренней самости и феноменального мира ощущений и чувств, метафизически личностного и свободы воли. В книге «Сознание объясненное» он выступил с провокативным заявлением, что «Все мы — зомби. Никто не является сознающим»<sup>1</sup>, а самость лучше всего толковать как «центр нарративной гравитации». Это заявление вызвало массу негодующих рецензий; по мнению многих критиков, книгу правильнее было бы назвать «Сознание игнорируемое» или «Самость исчезающая». Хотя есть и другие мнения: книга заставляет задуматься о глубине и масштабности головоломок феномена сознания и необходимости принципиально новых ключей для их разгадки.

Взамен картезианской «деформирующей» идеологии Деннет предлагает теорию, для которой он использует разные термины — «AI процессуально-информационная модель сознания», «функционалистский физикализм», «семиотический материализм», «материализм» и др. (Хотя в целом он скептически относится к определениям через «измы».)

В то же время — что существенно отличает Деннета от культур релятивистов, историцистов и постмодернистов — он консерватор в отношении ценностей науки, рациональности, истины и твердо отстаивает теоретическую значимость философского знания. Исходя, как он говорит, из «стандартного» представления о рациональности и науке, он занимает открыто материали-

<sup>1</sup> Dennett D. Consciousness Explained. P. 406.

8

стическую позицию. Он заявляет, что его «отправная точка — объективный, материалистический, рассматриваемый с позиции третьего лица мир физических наук» (1989, 5), а его философия является союзником и продолжением этих наук. Свою задачу он видит в том, чтобы посмотреть на сознание «взглядом третьего лица» через материалистическую перспективу науки, и готов доказать, что это — лучшая перспектива в современном понимании ментальных феноменов. Вместе с тем он спорит с Р. Рорти и другими постмодернистски настроенными коллегами, девальвирующими теоретический статус философии и утверждающими, что философские

рассуждения о сознании носят метафорический и эстетический характер. Не отрицая наличие в них эстетической компоненты и свободного проявления воображения, он по-другому видит взаимоотношение философии и науки. «Беда философии состоит в том, что она оказалась значительно более сложным феноменом, нежели это представляется ученому или человеку от искусства, поскольку она разделяет — и должна разделять— претензии и методы того и другого... Поэтому мой метод... очень серьезно принимает науку, однако тактика его более схожа с методами искусства»<sup>1</sup>.

Деннет — оптимист и сциентист и в этом отношении типично американский, верный духу Дьюи философ. «Загвоздка Вселенной» — тайна феномена сознания, — убежден он, вполне устранима при большей свободе воображения и мобилизации современной когнитивной культуры. Сегодняшний философ не может осмысливать сознание, не опираясь на то, что сделано в когнитивных науках, психологии, науках об искусственном интеллекте, в биологии и нейрофизиологии, и не выстраивая свои концепции как области философии науки. «Философия сознания... занимаясь концептуальными основаниями и проблемами наук о духе, превратилась в область философии науки. Это изменило форму и содержание философских теорий сознания, введя в дискуссии по традиционным

<sup>1</sup> Dennett D. *Elbow Room: the Varieties of Free Will Worth Wanting*. Oxford, 1984. P. 3.

9

вопросам много эмпирических данных и концептуальных инструментов новых научных подходов, а также постановку новых вопросов, возникающих из-за загадок и ловушек этих подходов»<sup>1</sup>. С другой стороны, подчеркивает он, когнитивные науки сегодня испытывают острую потребность в осмыслении их концептуального фундамента, что создает новую ситуацию востребованности философии.

В целом Деннет продолжил линии Дж. Райла на дискредитацию представления о сознании как «Духе в машине» и Л. Витгенштейна на развенчание идеи «приватного опыта». В чем-то он унаследовал и их логический бихевиоризм. Однако при подготовке диссертации его не удовлетворила царящая в Оксфорде атмосфера самоустраненности философии от науки. Им была высказана крамольная для лингвистического анализа мысль о необходимости натурализации философии сознания и применения к ней эволюционного принципа. Вспоминая свою учебу в Оксфорде, он говорит: «Свое самообразование я начал, всегда имея в виду не дающий покоя вопрос, как механические реакции «глупых» нейронов могут быть соединены в фабрику деятельности, умеющей на практике распознавать значения»<sup>2</sup>. Куайновские установки, усвоенные еще в студенческие годы в Гарварде, склонили его к выбору сциентистского панобъективистского вектора работы. В 60—70-х годах его воображение захватил начавшийся бум компьютеризации, идеи применения компьютерных возможностей в робототехнике, игре в шахматы, в особенности — открывающиеся перспективы исследования естественного интеллекта в сравнении с искусственным. Новаторская работа Герберта Саймона и Алана Ньюэла, в которой мыслительные действия человека сравнивались с операциями компьютера, пионерский шаг Хилари Патнэма, применившего идею «Машины Тьюринга» для объяснения сознания, утвердили его в мысли применения к объяснению сознания «универ-

<sup>1</sup> Dennett D. *Brainstorms*. P. XIV.

<sup>2</sup> Dennett D. *Self-Potrait*. Internet: <http://ase.tufts.edu/cogstud/pubpage.htm>. P. 1.

10

сальной машины Тьюринга». С годами у него стал нарастать, становясь доминирующим, мотив дарвиновского эволюционизма. Существенную роль в этом сыграло его знакомство с работами социобиологов, прежде всего Ричарда Докинса и Эдварда Уилсона, в которых биологическая и культурная эволюции представлены как сопряженный процесс. Влияния шли и с другой стороны. В этот период в англоязычной литературе активно обсуждались идеи «генетической инженерии» (не смешивать с «генной инженерией»), согласно которым предлагалось посмотреть на эволюцию как на своего рода «инженерное творчество» или «дизайнерскую работу» природы.

Какими бы ни были питающие корни философии Деннета, он создал собственную оригинальную

концепцию. Им было предложено много интересных концептуальных новаций — таксономия интенциональных установок, Модель Множественных Набросков, идея виртуальности машины сознания, гипотезы о соединении довербального и вербального, особая трактовка коэволюции биологического и социального и множество других неординарных гипотез о самосознающем себя интеллекте. Все эти концептуальные средства используются им для реализации извечной цели материалистов — обоснования единства мира и отсутствия пропасти между человеком и природой. Забегая вперед и оценивая его концепцию в целом, можно сказать, что метафизическая мечта Деннета состоит в том, чтобы показать, как объективно и без какого-либо агента в мире работают алгоритмы «Машины Дарвина», «Машины Тьюринга» и «Машины Сознания». Такое видение мира, считает он, вызревало давно, подходы к нему проложили выдающиеся умы. Юм был первым философом, попытавшимся в его ассоцианистской теории восприятия представить, как *мысли могут сами себя мыслить*. Мощный толчок к изображению хода объективных процессов без субъекта и сознания внес Дарвин, показав, что эволюционный алгоритм, проявляющийся в поразительном разнообразии природы, действует в виде слепого, *механического, бессознательного, не имеющего каких-либо мотивов* процесса. Вклад Тьюринга состоял в демонстрации воз-

11

можности изъять живого математика из мыслительных математических операций и передать эту работу машине, совершающей *мыслительные операции без мыслителя*.

Метафизическую мечту представить мышление без мыслителя, строго говоря, мог бы разделять и Гегель, живи он сегодня. Специфика Деннета состоит в подведении под нее натуралистической подкладки. Какими бы сложными ни были загадки сознания, языка и культуры, какие бы необычные вещи ни демонстрировал артефакт AI, все они в конечном счете — генетически связанные между собой продукты функционирования Матери-Природы, организации ее исходного (первичного) дизайна. Задача философа состоит в том, чтобы распознать и объяснить общие для природы, сознания и мира артефактов алгоритмы. Соответственно ответы на вопросы «как возможны сознание, самость, целеполагающее поведение?» и «как возможен AI?», напрямую связаны с ответом на вопрос «как возможна работа эволюции?». Он говорит: «Мы являемся роботами и мы созданы роботами — каждый из нас состоит из нескольких триллионов роботоподобных клеток, столь же лишенных сознания, как и составляющие их молекулы; тем не менее, совместно работая в гигантской команде, они производят всю без исключения деятельность, происходящую в сознающем агенте. Величайший вклад Тьюринга состоял в демонстрации возможности перевода кантовского вопроса в инженерный вопрос. Тьюринг показал нам, как можно поменять перволичностную перспективу Декарта и Канта на третьеличностную перспективу естественных наук и ответить на все вопросы — без философски значимого остатка»<sup>1</sup>. Деннет прошел школу аналитической философии и хорошо усвоил технику контекстуального анализа. На практике он выглядит, как мыслительные операции со значениями, показ того, что привычные словосочетания вроде «мой субъективный мир», «свобода воли» в разных контекстах и логически возможных мирах отнюдь не являются самоочевидными. Тем

<sup>1</sup> Dennett D. The Fantasy of First-Person Science // March, 2001. P. 1. Internet: <http://ase.tufts.edu/cogstud/pubpage.htm>.

12

не менее он не склонен преувеличивать значимость канонов логического и лингвистического анализа. Эти каноны позволяют яснее представить и более четко сформулировать проблемы, но вовсе не гарантируют их правильное решение.

Как и всякий аналитик, Деннет с большим недоверием относится к манере облекать требующие детального анализа проблемы в свободные одежды метафилософской риторики и обобщений (что свойственно, например, Ричарду Рорти). Не считает серьезными претензии некоторых философов дать окончательные ответы на фундаментальные философские вопросы. Он говорит: «Меня обычно не упрекают в излишней скромности; в рецензиях на мои книги частенько мелькают такие слова, как «высокомерие», «самонадеянность», однако в одном отношении я более скромнен, чем

это обычно свойственно философам. Я нерешителен и осторожен в высказывании окончательных выводов о Реальности, Истине, Значении, Времени, Каузальности и других великих сюжетах метафизики и эпистемологии... Я только взял на себя смелость разрабатывать некоторые наиболее поразительные выводы, вытекающие из стандартной научной картины мира, проще говоря, той картины, которая преподается в школе»<sup>1</sup>.

Свой стиль философствования он чаще всего называет «инженерным», или «упражнением в воображении», или «мысленными экспериментами». В заимствованный им из технического лексикона термин «инженерный» вкладывается прежде всего смысл «работы». Его интересуют организационные и функциональные принципы работы «машин» сознания, мозга, поведения, языка. Термин «инженерный» несет еще один смысл — серьезности. В его представлении философия — это тяжелый труд, не терпящий интеллектуальной лености, суждений обо всем на свете с высоты птичьего полета, желаний удивить экстравагантными идеями или литературными изысками. Иллюстрацией его убеждения в значимости тяжелой работы может служить одна из рассказанных им историй.

<sup>1</sup> Dennett D. Back from Drawing Board // Dennett and His Critics. Demistifying Mind. Ed. by B. Dahlbom. Oxford, Cambridge, 1993. P. 204.

13

Однажды жил-был парень, которому хотелось узнать, в чем состоит смысл жизни. Для этого он прошел тысячи миль и взобрался на вершину горы, где обитал мудрый гуру. «Скажи мне, в чем заключается смысл жизни?» — обратился он к нему.

«Скажу, конечно, — ответил гуру, — но для того, чтобы понять мой ответ, тебе нужно сначала создать теорию рекурсивных функций и математическую логику».

«Ты шутишь...»

«Совсем нет».

«Ну... тогда это неважно».

«Что ж, поступай как знаешь»<sup>1</sup>.

## 2. Презентация философии Деннета: трудности ее исследования

Наш рассказ о творчестве Деннета по своему жанру является историко-философским. Его практическая цель состоит в том, чтобы русскоязычный читатель получил представление о концепции сознания и самости Деннета — одной из самых ярких версий современного материализма. По сути, наше повествование сводится к трансляции культурной онтологии, в которой он работает, на понятный для российского читателя язык. В силу специфики российского философского образования студенты весьма скудно знакомы с англоязычной философией и ее стилистикой, чем, вероятно, объясняется отсутствие у них большого интереса к ней. Часто их отпугивают существующие здесь каноны строгости рассуждения. Сказывается и отсутствие у нас достаточного количества аналитической и комментаторской литературы, специально посвященной современным концепциям сознания<sup>2</sup>. Другая, связанная с ней прак-

<sup>1</sup> Ibid. P. 5.

<sup>2</sup> В 20—50-е годы проблемами сознания и мышления в Советском Союзе занимались главным образом психологи — Л.С. Выгодский.

14

тическая цель рассказа о философии Деннета состоит в том, чтобы развеять бытующие в российской философской литературе представления о «немодности» философской проблемы сознания, «устарелости основного вопроса философии», о том, что дилемму материализм-идеализм нужно сдать в архив<sup>1</sup>. Нам С.Л. Рубинштейн, А.Н. Леонтьев, В.В. Давыдов и др., развивавшие (в разном толковании) деятельностный подход к сознанию. (О судьбе деятельностного подхода см.: Лекторский В.А. Деятельностный подход: смерть или возрождение // Вопросы философии. 2001. № 2.) Среди философов, которые внимательно следили за тем, что происходит в мировой литературе по философии сознания, и осмыслили сам феномен сознания, следует назвать Д.И. Дубровского, Э.В. Ильенкова, Ф.Т. Михайлова, В.А. Лекторского, И.П.

Меркулова, С.Д. Балмаеву и др. Большую переводческую (и авторскую) работу проделал А.Ф. Грязнов, сделав доступным для русскоязычного читателя многие ключевые работы в этой области.

<sup>1</sup> Один мой коллега по Институту философии, когда зашел разговор об основном вопросе философии, который мы, естественно, учили «по Ленину», выразил глубокое убеждение, что нынешним российским философам он неинтересен. Такое заявление меня удивило: можно отбросить наивности марксистской аргументации в пользу теории отражения и вульгарный материализм в подходе к проблеме сознания (по сути являвшийся защитой дуализма), но сами вопросы отнюдь не канули в вечность. Они как составляли, так и будут составлять «твердое ядро» философии. Другое дело, что наши философы почему-то избегают работать в сфере «твердого ядра», предпочитая производные, периферийные, культурологические, а подчас и околофилософские темы. Показательно также, что во многих учебниках по философии, выпущенных в России в 90-е годы, тема сознания блистательно выпадает. Существует чисто фактический (библиографически достоверный) аргумент в пользу утверждения, что основной вопрос философии не потерял и, по-видимому, не потеряет, по крайней мере в обозримом будущем, своей значимости — это публикация (начиная с 50-х годов) огромного количества литературы за рубежом по проблеме сознания и реализма. Поток литературы не уменьшается по сей день.

15

хотелось бы, чтобы чтение этой книги убедило молодого читателя, что накал дискуссий вокруг феномена сознания в англоязычной философии не есть какая-то ее культурная особенность и не объясняется только реакцией на появление феномена искусственного интеллекта и более тонкие исследования нейрофизиологии мозга. С нашей точки зрения, повышенное к ней внимание в первую очередь объясняется тем непреложным фактом, что проблема сознания в совокупности с проблемой реализма является *фундаментальной, входящей в «твердое ядро» философии проблемой*. И точно так же, как это происходит в науке, характер разработки фундаментальных областей определяет характер и качественный уровень прикладных областей. Нам представляется, что российские авторы, рассуждающие о мышлении, о человеке, о коэволюции биологического и социального, абстрагируясь от понимания главного — сознания, — поступают легковесно.

С самого начала мы хотели бы предупредить читателя, что не ставим перед собой задачи противопоставить концепции сознания Деннета собственную альтернативную позицию: формулировка последней требует теоретико-проблемной работы, а это совсем иной жанр, предполагающий другую квалификацию и владение другим характером информации (сегодня нельзя, например, рассуждать о сознании и мышлении, не задумываясь о том, как работает искусственный интеллект и какие между ними аналогии и дизанalogии). Иначе говоря, наша работа ограничивается *дескриптивным анализом*. (Этим объясняется обильное цитирование источников.) В рамках этого жанра мы, разумеется, занимаемся реконструкцией взглядов Деннета. Мы постарались организовать текстовый материал таким образом, чтобы реализовать следующие задачи: 1) очертить общее проблемное поле, которое исследуется Деннетом (сознание, духовное-телесное, самость, опыт, феноменальное, интенциональное, интроспекция, свобода воли, творчество, личностное, язык, детерминизм и ответственность, естественный и искусственный интеллект и

16

др.); 2) обрисовать теоретический и исторический контекст, в котором он работает; 3) определить специфику его стратегии — деконструктивной и конструктивной ее сторон; 4) показать применяемую им когнитивную культуру; 5) выделить основные узлы, являющиеся мишенью критики, и в связи с этим обозначить основные оппозиции, сложившиеся в англоязычной философии сознания; 6) и, наконец, поскольку автора данной книги всегда интересовали метафизические аспекты философских конструкций, в особенности выявление имплицитной метафизики в декларациях антиметафизиков<sup>1</sup>, в настоящей работе этот интерес проявился в том, что мы постарались проследить, как исходная «минимальная» метафизика Деннета с годами перерастает в «максимальную» метафизику.

Одна из задач нашей книги — показать стилистику работы «философской лаборатории» Деннета, характер используемой им информации, разнообразие применяемых им инструментов, а также



этические резоны и личные амбиции, стимулирующие его в высшей степени трудоемкой работе. Одновременно нам хотелось показать, что важно, его реакцию на критику.

Более широкая цель книги состоит в том, чтобы на примере творчества Деннета «схватить» особенности *философской ментальности* времени, ибо, независимо от конечных результатов его материалистического проекта, он, безусловно, является «знаковой» фигурой англоязычной философии второй половины XX века. «Знаковость» сказалась в том, что в его творчестве нашли отражение такие ее характерные черты, как тенденция к соединению лингвистической традиции с натуралистической, отказ от упрощенного эмпиризма неопозитивистов и обращение к эволюционному когнитивизму, усиление связей философии сознания с когнитивными науками, интенсивное использование компьютерных аналогий и метафор, утверждение теоретического статуса философских про-

<sup>1</sup> Юлина Н.С. Проблема метафизики в американской философии XX века. М., 1978.

17

блем. В связи с последней особенностью заметим, что, независимо от активности кампании постмодернистов, девальвирующих теоретический статус философии и превращающих ее в «жанр литературы», магистральное движение нынешней англоязычной философии все же представлено натурализмом, гносеологизмом, объективизмом. Такими именами, как Д. Деннет, Дж. Сёрль, Дж. Фодор, П. Черчленд, Н. Блок, Д. Чэлмерс и многими другими, для которых философия является не прихотливым литературным дискурсом, а доказательным знанием.

Для российского историка философии творчество Деннета — сложный и трудоемкий объект исследования. Жанр аналитической философии вообще трудно поддается дескриптивному анализу (этим, наверное, объясняется малочисленность за рубежом описательных работ в сравнении с огромным количеством теоретико-проблемных работ). В англоязычной философии бытует устойчивое мнение, что аналитическую философию следует *показывать*, а не *описывать*. Философствовать — значит знать, как надо решать проблемы и демонстрировать профессиональное умение это делать, а не просто рассказывать о том, что делает другой человек. Отсюда сложилась практика разводить по разным отсекам и наделять неодинаковыми рангами философско-теоретические (проблемные) работы, с одной стороны, и метафилософские и историко-философские исследования, с другой. Сказанное во многом объясняет факт, что хотя Деннет не испытывает недостатка в комментариях, большинство из них — а это преимущественно работы философов-аналитиков и когнитивных ученых — носит проблемный и дискуссионный характер. Для многих авторов провокативная концепция Деннета (как в свое время концепция Витгенштейна) служит своего рода оселком, на котором они оттачивают свои собственные аргументы. Об этом свидетельствует, в частности, характер статей в сборниках, специально посвященных обсуждению его концепции: «Деннет и его критики: сознание без мистификации» (1993); «"Сознание объясненное" Дэниела Деннета» (1993); «Философия

18

Деннета: общая оценка» (2000); «Дэниел Деннет» (2002)<sup>1</sup>. Дескриптивных работ историко-философского жанра нам не встретилось. (Возможно, из-за скудости новых поступлений литературы в московские библиотеки.) Одним словом, наша работа в известной мере восполняет этот пробел.

Серьезную трудность историк испытывает из-за того, что Деннет стремится одной ногой стоять в философии, а другой — в науке. Интересуясь (подчас профессионально) тем, что происходит в когнитивных науках, биологии, теориях искусственного интеллекта, он пишет специальные работы и часто вовлекается в дискуссии по поводу толкования тех или иных конкретных эмпирических фактов или выводов. (Это касается исследований по искусственному интеллекту, толкования эволюционной теории и др.) Не являясь специалистами в этих областях, мы не будем особо вдаваться в эту сторону его деятельности. Хотя, конечно, нас интересуют общеполитические выводы, которые им делаются из осмысления наук, и та философская критика, которой он подвергает те или иные научные позиции. Одна из целей нашей книги состоит в показе на материале философии Деннета, что «натуралистический крен», характерный для многих



англоязычных теорий сознания, сопряжен с выходом за чисто философские концептуальные схемы к осмыслению тех или иных эмпирических фактов.

Особая сложность экспликации его взглядов связана с интенсивным применением им принципа эволюционизма. Продолжая «дарвиновскую» натуралистическую традицию американской мысли, он развивает организмический, процессуальный взгляд на мир и стремится подчеркнуть транзитив-

<sup>1</sup> Thought and Object: Essays on Intentionality. Ed. by Woodfield, Oxford, 1982; Daniel Dennett's «Consciousness Explained». Symposium // Inquiry. V. 36. № 1/2. 1993; Dennett and His Critics: Demystifying Mind. Ed. by B. Dahlbom. Oxford, Cambridge, 1993; Dennett's Philosophy: A Comprehensive Assessment. Ed. by D. Ross, A. Brook, D. Thomson. Cambridge MA, 2000; Daniel Dennett. Ed. by A. Brook and D. Ross. Cambridge, 2002.

19

ность разных явлений. Отсюда у него нелюбовь к жестким демаркациям, а вместе с этим — и к четким формулировкам. На одном из симпозиумов, отвечая критикам, он сказал, что предпочитает работать «на срезах», «стыках» или «на краях» («Living on the edge»). Реально это означает выбор в качестве предмета пространств стыковок и переливов биологического и социального, живого и неживого, естественного и искусственного, сознательного и бессознательного, науки и философии. Фокусировка внимания на транзитивности сказывается не только в избегании жестких формулировок: он предупреждает, чтобы в его рассуждениях не искали цельную систематическую доктрину. По его словам, «результатом его изысканий явилась теория, выглядящая неуклюжим и неэлегантным гибридом, неименоваемой мешаниной из частей теории, если сопоставлять ее с традиционным паттерном категорий. Поскольку, как я полагаю, моя теория обрисовывает природу на ее стыках, тем не менее я склонен утверждать, что именно традиционный паттерн был сломан. По этой причине я до сих пор воздерживался от именованной моей теории...»<sup>1</sup>. Само собой разумеется, что его нелюбовь к систематизации существенно затрудняет описательные задачи исследователей его творчества.

Не сразу можно постичь метод аргументации Деннета. Он состоит в проведении мысленных экспериментов, рассказывании историй и приведении примеров. Взятые из обыденной жизни (или логически возможных виртуальных ситуаций), истории кажутся простыми и тривиальными. Чтобы понять их аргументативную цель, от русскоязычного читателя требуются, помимо способности вычитывать философские смыслы в иноязычных текстах, трудоемкие усилия по вхождению в культурную онтологию английского языка, в частности — в семантические пласты метафорического и идиоматического языков.

Серьезно осложняет задачу историка то обстоятельство, что по натуре Деннет полемист, а его книги — реакции на дис-

<sup>1</sup> Dennett D. Brainstorms. P. XIV.

20

куссии с многочисленными оппонентами и коллегами. Чтобы понять его позицию, нужно понять позиции его визави. Список авторов, с которыми он ведет интеллектуальную полемику, огромен: Б. Скиннер и У. Куайн, Л. Витгенштейн и Дж. Райл, Н. Хомский и К. Поппер, У. Селларс и Х. Патнэм, М. Арбиб и Дж. Фодор, Дж. Сёрль и Т. Нагель, Н. Блок и С. Кларк, К. МакГинн и Р. Рорти, Д. Чэлмерс и Р. Пенроуз, С. Гоулд и П. Черчленд и многие, многие другие. Разумеется, изобразить длящийся почти полвека полилог практически невозможно — слишком велик круг людей, участвующих в нем, и слишком многообразны их взгляды. Для этого нужно писать специальную работу. Чтобы читатель в какой-то мере почувствовал роль и значимость дискуссионной атмосферы, в которой «варилась» и «созревала» его философия, вторая часть книги специально посвящена теме «Деннет и его критики». Она состоит из разделов: «Деннет versus Нагель», «Деннет versus Сёрль», «Деннет versus Рорти», «Деннет versus Поппер». Учитывая историко-философский жанр книги, предполагающий типологизацию и изображение картины философской жизни широкими мазками, автор данной работы довольно часто прибегает к «измам». Историки всегда знали, что классификация с помощью «измов», полезная для целей презентации и преподавания, огрубляет и схематизирует реальное положение дел. Сегодня

особенно нужна осторожность в их применении. Почти все фигурирующие в философии сознания классификационные термины — «материализм», «физикализм», «дуализм», «функционализм» «информационная теория», «эмерджентизм», «элиминативизм», «редукционизм», «когнитивизм» — у разных авторов имеют разное толкование. Например, приверженцы одного и того же лагеря физикализма, сходясь в отрицании дуализма, по-разному толкуют термины «ментальное», «феноменальное», «физическое», «физика», «каузальность», «объективность». Столь же многообразно толкуется смысл «функционализма». И еще одно немаловажное наблюдение. Современных философов сознания в первую очередь занимает существо дела — представить убедительные,

21

«разящие» аргументы в пользу тех или иных тезисов; их мало беспокоят «лагерные» ярлыки и направленческая чистота<sup>1</sup>. К слову сказать, их мало волнует и метафилософская проблематика — чем должна, а чем не должна заниматься философия. И даже разговоры об эпистемологических инструментах потеряли свою приоритетность по сравнению с главным — объяснением факта сознания.

И последнее, о чем следует сказать, хотя на самом деле это первая трудность, с которой сталкивается исследователь творчества Деннета, — перевод ключевых терминов. В первую очередь это относится к «ментальным» понятиям и связанным с ними терминам (*mind*, *consciousness*, *belief*, *awareness*, *sense*, *sentience*, *qualia*, *introspection*, *mental representation*, *disposition*, *intentional stance*, *mind-body problem*, *intuition pumps*, *propositional attitude*, *experience*, *content*, *supervenience* и др.), которые не имеют точного аналога в русском языке. Лингвистические неудобства возникают и с переводом компьютерных терминов: *software*, *hardware*, *information*, *virtual machine*, *input*, *output*, *drive* и др. Многие из этих терминов мы даем в английском написании по причине того, что у Деннета они имеют философский, а не точный технический смысл, к тому же в переводном виде их трудно ввести в строй и стилистику русского языка<sup>2</sup>.

<sup>1</sup> Проф. Роберт Кирк, активно участвующий в баталиях философов сознания, по этому поводу заметил, что вопрос «дуализм versus физикализм» «безотносителен к нашим проблемам» (Kirk R. *Raw Feeling. A Philosophical Account of the Essence of Consciousness*. Oxford. Clarendon Press, 1996. P. 7).

<sup>2</sup> Работающие с англоязычной литературой по философии сознания прежде всего сталкиваются с трудностью перевода ключевых терминов «*mind*» и «*consciousness*». Слово «*mind*» в английском языке часто, хотя далеко не всегда, используется с психологическим оттенком. К нему прибегают, например, при обсуждении психофизиологической проблемы (*mind-body problem*) и когда речь идет о психике животного. Оно используется для обозначения различных форм

22

### 3. Тайна феномена сознания

История дискуссий по проблеме сознания в англоязычной философии за последние 50 лет является собой один из примеров интеллектуальной драмы или, прибегая к образу из другого жанра литературы, интеллектуального детектива. Как и положено для этого жанра, вначале имеет место «тайна» («преступление»), затем идет расследование. В ходе расследования каждый детектив применяет свою особую стратегию, убежденный, что только она приведет к обнаружению «преступника». Если пойти в глубь веков, первым «детективом», обнаружившим «тайну сознания», был Платон, заметивший несовпадение идеальных форм содержания сознания с миром повседневных явлений. Декарт обратил внимание на другую сторону тайны — принципиальное отличие *res cogitans* от *res extensa* активности, обычно относимой к «ментальной», — чувств, эмоций, настроений и др. Некоторые переводят «*mind*» и «*consciousness*» словами «психика» и «сознание», однако это не годится по отношению к философам, которые не проводят между ними различия. Слово «*consciousness*» чаще всего, хотя тоже не всегда, используется для обозначения когнитивного содержания сознания (в противоположность феноменальному и подсознательному). В последнее время эти различия не очень заметны, оба термина используются достаточно свободно и часто перекрывают друг друга. В книгах с названиями «*Philosophy of mind*» и

«Philosophy of consciousness» нередко обсуждаются примерно одни и те же проблемы. У Деннета эти два понятия взаимозаменяемы, хотя он отдает предпочтение более когнитивному термину «consciousness». Неясности существуют и в отношении других ментальных терминов: термин «awareness», не имеющий аналога в русском языке, мы переводим как «находиться в состоянии осознания» или «осознание»; слово «sentience» мы переводим как «ощущение», «чувство». Одним словом, читатель должен учитывать, что смыслы употребляемых «ментальных» понятий зависят от интерпретации тем или иным автором феномена сознания.

23

и в то же время их интеракцию в мыслительном акте. Английские эмпирики увидели много загадочного в феноменальном опыте: Локк занялся исследованием природы вторичных качеств, а Юм бился над причинами несовпадения чувственных впечатлений и объектов, которые их вызывают. Brentano подметил в тайне аспект, который не видели «детективы» до него: ментальные понятия, в отличие от понятий, относящихся к физическим вещам, имеют интенциональную нагруженность. К XX веку проблема сознания, которая Декарту казалась рядовой и вполне разрешимой, обросла частоколом других тайн и превратилась в *крепость*. Конечно, появилась новая когорта детективов, которые усомнились в свидетельских показаниях, добытых их предшественниками, а тем более в убедительности их версий раскрытия тайны сознания. Владея более совершенными когнитивными инструментами и вдохновляемые оптимистической верой в способность человека познать самого себя, они с разных сторон ринулись на штурм крепости, именуемой «Тайна Сознания». И только после многочисленных неудач в объяснении того, как идеальные формы мысли соотносятся с видимым миром, как *res cogitans* сцепляется с *res extensa*, как ощущения дают представление о внешнем мире, тайна сознания была переведена в разряд *особой*. Она обрела статус «Великой Тайны» или, как говорил Шопенгауэр, «Загвоздки Вселенной». Этот статус она сохранила по сей день.

Конечно, существует множество других великих тайн. Например, нам многое неизвестно о происхождении Вселенной, строении материи, природе времени, пространства, гравитации, возникновении жизни и репродукции. Однако в отношении этих тайн по меньшей мере есть согласие, где и как искать их разгадку; считается, что все они в принципе подвластны научному, рациональному изысканию. Что касается тайны сознания, то здесь царит замешательство: неясно, с какой стороны к ней можно подходить и какими средствами можно раскрыть. Должна ли этим заниматься философия, или когнитивная психология, или биология, или физика, или ка-

24

кая-либо новая междисциплинарная отрасль знания? Или отбросить науку и довериться собственной интуиции и здравому смыслу? (Разумеется, что для академических философов, работающих в престижных университетах, концепции которых мы будем разбирать, апелляция к внерациональным и мистическим средствам неприемлема.)

В чем специфика тайны сознания? В обычной жизни, как правило, люди не обременяют себя вопросом «что такое сознание?», считая его данной от Бога или Природы способностью ориентировать нас в мире, подобно той, которой обладает сердце для перекачки крови в организме. Когда же они задаются им, ответ на него имеет одну примечательную особенность — асимметрию интуитивного и теоретического представления о сознании. С одной стороны, когда мы обращаем вопрос к самим себе, ответ на него кажется предельно ясным, ибо для каждого человека нет ничего более близкого, интимного и достоверного, нежели собственное сознание, внутреннее Я (самость). В отличие от внешних объектов, знание о существовании которых может быть ошибочным и релятивным, человек абсолютно уверен в существовании феномена его собственного сознания; здесь ошибки невозможны, поскольку оно дано ему непосредственно, а не опосредованно, и удостоверяется самим фактом осознанности. То есть все мы в повседневности, или, как сейчас говорят, в фолк-психологии, готовы подписаться под формулой — *cogito ergo sum*. С другой стороны, когда философы или психологи пытаются дать теоретическое определение сознанию (объективно или с позиции третьего лица) или хотя бы обозначить составляющие его элементы и приблизительные границы, феномен расплывается, теряет очертания, ускользает от

«определения в понятиях». (В современной философии это сказывается в разногласии в понимании того, что следует считать состояниями сознания, данными опыта, восприятиями, чувственным, когнитивным, сознательным, бессознательным и др.) Все прошлые и нынешние попытки нарисовать более или менее непротиворечивую теоретическую картину сознания сталкиваются с тем, что перед

25

исследователем предстает запутанный клубок феноменов, именуемых «сознанием», распутать который оказывается в высшей степени трудоемкой задачей с весьма туманными перспективами. Практически все современные философы сходятся в выводе, что сознанию невозможно дать *логическое* определение с приведением необходимых и достаточных условий. Надеяться, что теория сознания может быть выстроена в виде когерентной системы принципов, — напрасное ожидание. Таким же самообманом будет поиск определения сознания на основе элементарных, или базисных, структур, ибо в акте сознания задействовано такое количество разнородных процессов, редуцировать которые к чему-то элементарному в принципе невозможно.

Ускользаемость сознания от определений еще у Юма вызвала подозрение: «А не является ли оно фикцией?» В известной статье «Существует ли сознание?» Уильям Джемс открыто сформулировал это сомнение. У ряда сегодняшних философов (элиминативистов) это подозрение вылилось в радикальный диагноз об иллюзорности этого феномена. Они говорят, что как такового его нет в природе, что теоретически правомерно говорить о сознании только на основе внешне фиксируемых проявлений — на производимых текстах, нейрофизиологических процессах, физическом поведении, которые фиксируются intersubjectivными методами. Очень многие философы не согласны с таким радикализмом. Они убеждены, что сознание, хотя и не поддается анализу объективными методами, существует как реальность. Просто оно оказалось значительно более сложным явлением, чем представлялось ранее. Вполне рационалистически и даже сциентистски мыслящие авторы говорят о его принципиальной неопределяемости, о его «мистике» и «вечной загадке». И о кризисе научной объективности, коли с помощью рациональных методов невозможно объяснить этот феномен (Дж. Сёрль, К. МакГинн и другие). Трудность нахождения общезначимого определения сознания нашла отражение в зарубежных словарях и энциклопеди-

26

ях. В «Энциклопедическом словаре по психологии» (1983) говорится, что этот термин используется весьма свободно, разные авторы фиксируют разные состояния сознания, и ни одно из приводимых определений нельзя считать непротиворечивым<sup>1</sup>. В более позднем «Международном словаре по психологии» (1989) об определении сознания говорится с безнадежностью: «Сознание — завораживающий, но призрачный феномен: невозможно специфицировать, что же оно такое, что оно делает, и почему оно эволюционировало. Ничего такого, что стоило бы читать, о нем не написано»<sup>2</sup>. Несмотря на призрачность феномена сознания, он продолжает завораживать, и философы не перестают искать ключи к его разгадке.

Люди не могут не любопытствовать, как так получается, что немыслящая материя в состоянии мыслить? В самопонимании обычного человека не укладывается, как может мозг — комок биологического вещества — быть вместилищем полета мысли, свободы воли, самосознания, личностных переживаний, моральности и ответственности, одним словом, всего того, что составляет человеческое бытие. Ведь всего этого нет в биологическом и физическом мире. Или, если сформулировать недоумение в другой перспективе, как может нечто неосознаваемое, именуемое нами «сознанием», обладать свойствами, которые выходят за пределы закономерностей биологической природы, и создавать артефакты — мир культуры? Как не имеющие пространственно-временного измерения мысли могут взаимодействовать с телом, подчиняющимся законам физики?

Обычная интуиция о самом себе толкает человека к естественному дуализму: мозг — это одно, а сознание — совершенно другое, т.е. к картезианству здравого смысла. Опираясь на эту интуицию, философы выстраивали и продолжают строить всевозможные софистичные версии удвоенного

мира. Это

<sup>1</sup> The Encyclopedic Dictionary of Psychology. Ed. by R. Harre and R. Lamb. Oxford, 1983. P. 114.

<sup>2</sup> Цит по: Chalmers D.J. The Conscious Mind. In Search of Fundamental Theory. N.Y., Oxford et al., 1996. P. 3.

27

одна сторона истории. Другая сторона истории состоит в том, что дуализм, если не принимать его религиозную версию, чреват бесконечным регрессом, ибо оставляет без ответа вопросы: «откуда» и «почему» *res cogitans*. Рационалистически и критически мыслящие люди никогда не мирились с объяснениями на основе бесконечного регресса. Поскольку поступь науки уверенно направлялась в сторону накопления твердо установленных фактов и все большей объективности и отстраненности от субъективно-личностного, у исследователей возникло другое, не менее сильное интуитивное чувство, а именно, что тайну сознания поможет раскрыть накопление конкретных, intersubjectively проверяемых фактов о мозге, о деятельности когнитивного аппарата, о действующих здесь причинно-следственных связях. И что на их основе можно будет сказать что-либо фундаментальное о сознании. Эти чувства служили и служат внутренним мотивом различных течений научно ориентированного материализма.

В XX веке, в особенности во второй его половине, фактов о работе мозга накопилось много. Существенный прогресс наметился с компьютеризацией эмпирической психологии и появлением когнитивных наук с их междисциплинарными подходами. Много эмпирических данных получено, например, о визуальном восприятии, памяти, внимании, воображении, овладении языком, обучении, принятии решений и др.<sup>1</sup> Тем не менее тайна сознания осталась тайной. Сегодня, например, известно, что человеческий мозг представляет собой структуру огромной анатомической и физиологической сложности и имеет (по приблизительным подсчетам ученых)  $10^{12}$  нейронов и  $10^{15}$  связей между ними. И все же увеличивающийся массив знания о структуре и деятельности мозга не снял естественное интуитивное недоумение, как может чрезвычайно сложная, но

<sup>1</sup> Об открытиях в современных науках — генетике, нейрофизиологии, нейропсихологии, когнитивной психологии, компьютерной и др. — см.: Меркулов И.П. Информационная природа сознания // Полигносис, 2000. № 4.

28

все же осязаемая биологическая материя вмещать в себя еще что-то добавочное, что не является материей, — мысли. Ученые верят, что будут открыты уникальные функциональные свойства человеческого мозга, объясняющие, скажем, способности к саморефлексии или усвоению языка. Однако большой уверенности, что эти открытия снимут ореол тайны с сознания, нет. Многие сомневаются в возможности получения вразумительного ответа на вопрос, как могут знания о фактах функционирования нервной системы, какими бы они ни были полными и всесторонними, объяснить, скажем, уникальное восприятие индивидом хорошей музыки, пейзажа или поэмы?

Одним словом, без большой натяжки можно констатировать, что, несмотря на безусловные успехи в интерпретации ряда важных сторон сознания, ни у ученых, ни у философов на сегодня нет ясности, что следует считать фактом сознания, независимо от тех или иных теорий сознания. Именно эта напряженность между нашей неспособностью представить, как материальный мозг может быть вместителем мыслей, субъективных ощущений, эмоций, и уверенностью в том, что раскрытие каузальных связей и накопление фактов о каком-либо объекте дает нам объективное понимание этого объекта, является источником оппозиций в философии сознания и реанимации противостояний панобъективизма и субъективизма, монизма и дуализма. Этот источник не иссяк и по сей день. За новыми техническими терминами и жаргонами, обозначающими возможные позиции в логическом пространстве философии сознания, скрывается примерно старый расклад сил. Кто-то склоняется к пессимистическому убеждению, что тайна сознания навеки останется тайной, а кто-то оптимистически верит, что она может быть раскрыта.

Деннет находится в лагере оптимистов. Назвав свою книгу «Сознание объясненное», он, конечно, бросил вызов распространенному мнению о сокрытости тайны сознания от нашего разума. По его



признанию, в его проекте материалистического объяснения сознания путем его демистификации многие люди увидели «акт интеллектуального вандализма и

29

оскорбление последней святыни человечества»<sup>1</sup>. Они нашли название книги претенциозным, а автора — чересчур самонадеянным перед величием тайны сознания. Говорят, что автор предрекает период «пост-сознания», когда сознание «перескочит к науке», а понятия морального агента и свободы воли, любви и боли будут сброшены в архив.

На самом деле, говорит Деннет, все это не так. Просто на тайну сознания следует посмотреть по-другому, примерно так, как мы смотрим на тайну любви или денег. «Сознание, как и любовь или деньги, является феноменом, который в значительно большей мере зависит от ассоциирующихся с ним понятий. Хотя, как и любовь, оно имеет развитую биологическую основу, подобно деньгам некоторые из его наиболее существенных характеристик выросли внутри культуры, а не просто каким-то образом воплощены в физических структурах»<sup>2</sup>. Чтобы объяснить сознание, нужно перевести разговор о нем в сферу культуры, а порожденные декартовской традицией иллюзии отнести к разряду исторических наивностей. Поступить так, как поступила современная космология, которая перевела представления древних греков о небесных светилах, как о движущихся по небу колесницах богов, в разряд чего-то вроде комиксов, или что сделала биологическая теория ДНК, элиминировав популярную когда-то идею *elan vitae*. Потеря невинности в отношении сознания будет компенсирована лучшим пониманием очень многих вещей — научного, социального, теоретического и морального порядка. Сказанное не означает, что с тайной сознания вообще будет покончено. История научных открытий говорит о том, что раскрытие одних тайн порождает более интересные тайны. Новый взгляд на феномен сознания, возможно, породит более интригующие загадки. Во всяком случае, история науки свидетельствует, что загадки, не дающие покоя современным космологам и биологам, более масштабны, нежели тайны, мучившие ученых прошлого.

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 21.

<sup>2</sup> *Ibid.* P. 24.

30

#### 4. История атак на дуализм

Прежде чем перейти к рассмотрению содержательных блоков концепции сознания Деннета, имеет смысл сделать небольшой обзор интеллектуальной истории, на протяжении которой формировались и развивались его взгляды.

В общей форме можно сказать, что по своим метафизическим установкам Деннет физикалист. Не в том смысле, что он занимается философией физики (как раз физикой он не занимается: его научные интересы лежат в области биологии, когнитивной психологии, наук об искусственном интеллекте), а в смысле убеждения, что сознание следует объяснять только объективными, каузальными, не противоречащими физическим законам интересубъективно удостоверяемыми факторами. Поэтому его концепцию лучше всего идентифицировать на фоне полувековой истории современных вариантов физикализма.

История физикализма (или современного материализма) достаточно драматична. Ее истоки кроются в оптимистической вере Декарта в способность рационального человека познать самого себя. В то же время составляющие ее события представляют собой серию атак на картезианский дуализм — «последний бастион» традиционной философии, который представляется вызовом рационализму. Драматизм усугубляется тем фактом, что, несмотря на непрерывные атаки с использованием более совершенной техники наших дней, бастион сознания так и не пал, а число его защитников не сокращается. Не потому, что находящиеся в нем адепты дуализма, ментализма, эпифеноменализма и др. для отражения атак вооружены более изощренным оружием, а потому, что на стороне защитников крепости — «простой народ», или то, что называют «фолк-психологией»; проще говоря, психологическое убеждение обычного человека, что его тело — одно, а мысли — совершенно другое.



Атаки на бастион сознания физикалисты начали с демистификации ментального языка, на котором мы обычно гово-

31

рим о наших психологических состояниях. Среди первых, кто задал импульс физикалистскому движению в философии сознания, были Рудольф Карнап и Карл Гемпель. Строго говоря, для них главным было не разгадка тайны сознания, а обоснование возможности единой науки и знания. В рамках этой стратегии сознание выглядело аномалией, которую нужно было либо объяснить на основе объективистской модели, либо вообще устранить. Еще в работах 30-х годов и прежде всего в статье «Психология на физическом языке» (1932—1933)<sup>1</sup> Карнап сформулировал тезис физикализма, согласно которому так называемые психологические высказывания подлежат трансляции (переводу) на физический язык. Более точно, каждое психологическое высказывание следует соотносить с физическими событиями, происходящими в теле той или иной личности. Такая трансляция делает психологию частью единой науки, основанной на физике. Гемпель подчеркивал, что психологические высказывания не являются бессмысленными и в принципе они верифицируемы, но только в случае перевода их в предложения, в которых отсутствуют ментальные категории психологии и содержатся только категории физики. Проект трансляции неопозитивистов не был реализован по разным причинам. Прежде всего потому, что он требовал доказательства возможности строгой редукции языка психологии к языку нейрофизиологии, а затем языка биологии к языку физики. Возможность такой редукции никем не была показана. А также потому, что проект трансляции во многом был «завязан» на бихевиористском проекте Берреса Скиннера. Скиннер считал, что можно полностью объективировать психологический процесс, ограничившись описанием того, как психологические акты функционально и операционально, то есть на основе установления причинно-следственных связей, проявляются в поведенческих актах. Когнитивный процесс у него сводился к схеме «стимулы—черный ящик—реакции», в которой менталь-

<sup>1</sup> Carnap R. *Psychology in Physical Language // Logical Positivism*. Ed. A.J. Ayer. N.Y., 1959. Ориг. публ.: *Erkenntnis*. V. III, 1932/1934.

32

ное сознание было излишним. В дальнейшем тактика бихевиоризма заменить сознание «черным ящиком» была подвергнута серьезной критике, а схема в целом была признана неудовлетворительной.

Несмотря на критику тактики неопозитивистов и бихевиористов, их стратегические идеи не были преданы забвению. Идея Карнапа и Гемпеля о единой науке продолжает оставаться регулятивным идеалом практически во всех последующих физикалистских проектах. То же самое можно сказать о бихевиоризме. Бихевиористский подход, концентрирующий внимание на *формах поведения* в зависимости от получаемой информации, отнюдь не сошел со сцены; видоизменившись, он получил дальнейшее развитие (в частности в философии Деннета)<sup>1</sup>. Правда, говоря о бихевиоризме, следует учитывать, что в сравнении со скиннеровским толкованием понятие «поведение» было существенно расширено. Стало рассматриваться разное поведение — тела, процессов мозга, психологических состояний, речевой деятельности, лингвистических единиц, компьютерных программ, информационных систем и др.

Особое слово в объективации сознания сказал *логический бихевиоризм*, который обычно связывают с именами Л. Витгенштейна, Дж. Райла, Д. Льюиса. В отличие от психологического бихевиоризма, интерес которого был сосредоточен на психических реакциях организма, Райла интересовало прежде всего интеллектуальное поведение в форме семантического и грамматического поведения слов, которое проявляется во внешне наблюдаемом поведении. Разговор о сознании, который обычно ведется с использованием ментальных терминов, Райл предложил перевести в разговор о лингвистическом поведении. Или, в более слабом варианте, большую часть из того,

<sup>1</sup> Учитывая это, некоторые обозреватели не исключают, что со временем свободный от упрощенных и грубых толкований бихевиоризм «будет, наряду с логикой, рассматриваться как

главный оригинальный вклад в философию нашего нынешнего столетия» (Dahlbom Bo. Editor's Introduction // Dennett and His Critics. P. 4.

33

что в обычном языке классифицируется как разговор о сознании, корректнее объяснять в терминах, относящихся к фиксируемым высказываниям и наблюдаемым поступкам, а не в ментальных терминах, относящихся к данным интроспекции<sup>1</sup>. Веру в существование наряду с физическим миром особой ментальной реальности Райл связал с ошибками обычного словоупотребления. Его слова «история двух миров — это миф», его утверждение о неправомочности разговора о сознании как «Духе в Машине», или своего рода «гомункулусе», стали лозунгами в наступлении аналитической философии на дуализм и феноменализм (в частности, были положены Деннетом в основу стратегии объяснения сознания).

Л. Витгенштейн, как и Райл, не был физикалистом. Тем не менее его антикартезианство и антифеноменализм, его аргументы, направленные против существования особого приватного языка опыта, широко используются физикалистами. Согласно Витгенштейну, за ментальными терминами не стоят никакие сущности или референты; их значения зависят от употребления в определенном языке и от контекста конкретной «языковой игры». То есть они не относятся к «приватному» в смысле внутреннего интроспективного опыта, их значение выражено в поведении, которое можно фиксировать и наблюдать публично.

Важными следствиями импульсов, заданных неопозитивистами и философами лингвистического анализа, были следующие: философы заговорили о сознании с «лингвистическим акцентом»; философские проблемы стали пересматриваться через призму анализа языка, на котором мы говорим о них. Не поддающиеся такому анализу проблемы отбрасываются как неперспективные. Осмысленность и правомочность понятия «сознание» и всех прочих категорий ментального языка ставятся в прямую зависимость от приведения интерсубъективно признаваемых доказательств; данные интуиции и самоотчета субъекта в их число не входят. В более широкой перспективе

<sup>1</sup> Райл Дж. Понятие сознания. М., 2000. С. 334.

34

принятие в качестве признака сознания публично наблюдаемого лингвистического поведения означало *денатурализацию сознания и принятие социолингвистической парадигмы*: объяснение сознания целиком переводится в сферу социального и коммуникативного. Конечно, против социологизма и бихевиоризма высказывались и серьезные возражения. Одно из них принадлежало Ноаму Хомскому. Он высказал гипотезу о наличии у организмов внутреннего механизма или генетически унаследованной диспозиции к усвоению языка и грамматики. Исходя из нее, языковое и ментальное поведение невозможно объяснить без обращения к врожденной, то есть *небихевиористской*, компоненте<sup>1</sup>.

Говоря о развитии объективистской линии в философии сознания, нельзя пройти мимо *физикализма У. Куайна*. При выборе языка онтологии для Куайна принципиальными были два момента. Во-первых, установка на простоту: исходить из прагматических соображений и не умножать без надобности количество сущностей, с помощью которых мы описываем мир. Во-вторых, установка на холизм и онтологическую относительность теоретических схем. Все различаемые нами в мире сущности зависят от выбранного языка или теоретического каркаса; быть «чем-то» — значит иметь значение в рамках данной теории, в рамках другой теории это «что-то» может не иметь предиката бытия. Исходя из этих установок, Куайн заявил, что все фиксируемые нами в мире различия есть различия в позициях, состояниях и изменениях физических тел, поэтому говорить о существовании ментальных сущностей мы не вправе. В истории мысли были предложены три возможных уровня объяснения человеческого интеллектуального поведения: менталистский уровень, бихевиористский уровень (включая интеллектуальное лингвистическое поведение) и физиологический уровень. Менталистский уровень Куайн оценил как тупиковый и наименее удовлетворительный, а наиболее предпочтительным счел физиологический. Одна-

<sup>1</sup> Chomsky N. Language and Mind. N.Y., 1968.

35

ко, исходя из большей разработанности на сегодня языковой, а не физиологической теории, в качестве приемлемого варианта признал лингвистический бихевиоризм. (Согласившись вывести ментальность в языковую сферу, Куайн связал ее, прежде всего, с языком и социальным обучением.) Тем не менее он не считал лингвистический бихевиоризм оптимальной стратегией; в этом случае мы еще находимся на полпути от обскурантизма ментализма к будущему прогрессу нейрофизиологии<sup>1</sup>.

Интерес Витгенштейна и Райла был направлен прежде всего на анализ логики языка психологии и выявление ошибок способа говорения о ментальном, а не на конкретное объяснение того, каким образом наше сознание соотносится с мозгом. Оба они дистанцировались от происходящего в науке, считая, что научные открытия и методы не имеют значения для философски-концептуального решения проблем. В отличие от них Куайн сделал ставку на науку, точнее — на будущее развитие нейрофизиологии, однако он специально не разрабатывал тему «сознание/тело». На этом фоне новым словом в старом споре о сознании прозвучала *теория тождества*, сконцентрировавшая внимание на проблеме «сознание/тело» и апеллировавшая к науке. В 1958 году Герберт Фейгл, когда-то входивший в «Венский кружок», опубликовал в «Миннесотских исследованиях по философии науки» статью «"Ментальное" и "физическое"», а в следующем году австралийский философ Дж.Дж. Смарт выступил со статьей «Ощущения и процессы мозга»<sup>2</sup>. Независимо друг от друга Фейгл и Смарт предложили новый подход к проблеме сознания, который затем был назван «теорией тождества» (mind-body identity theory). Его другие на-

<sup>1</sup> Quine W.O. Facts of the Matter // Essays on the Philosophy of W.O. Quine. Ed. by. R. Shanan and Ch. Swoyer. Norman, 1979.

<sup>2</sup> Feigl H. The «Mental» and the «Physical» // Minnesota Studies in Philosophy of Science. Minneapolis, 1958. V. 3. P. 370-457; Smart J.J.C. Sensations and Brain Processes // Philosophical Review. LXVII. P. 141—156.

36

звания — «материализм центральных состояний» (central-state materialism), «теория состояний мозга» (the brain state theory). Смысл этого подхода состоял в том, чтобы показать, что хотя ментальные и физические термины семантически различаются между собой, они относятся к одним и тем же референтам. Термины «ментальное» и «физическое» обозначают одно и то же, подобно тому как термины «Утренняя звезда» и «Вечерняя звезда» имеют один и тот же референт — планету Венера. Работы Смarta и Фейгла (к ним можно добавить также работы английского психолога Ю. Плейса и австралийского философа Д. Армстронга) неожиданно для авторов получили широкий отклик, а теория тождества обрела многочисленных сторонников и противников. Историческая значимость статей этих авторов состояла в том, что основательно подзабытую «декартовскую» проблему духовного и телесного (в США ею всерьез интересовался У. Джемс) они снова ввели на авансцену философских споров. Теория тождества соответствовала оптимистическому настрою, возникшему в философской среде в связи с интенсивными научными исследованиями интеллектуальной деятельности. Поэтому вполне естественным казался вопрос: а почему ментальность не может оказаться просто процессами мозга, подобно тому как свет оказался электромагнитным излучением, а гены — молекулой ДНК?

Теория тождества пыталась ответить на два вопроса. На вопрос «что такое ментальные события?» она отвечала: каждое ментальное событие идентично физическому событию мозга, а на вопрос «что общее имеют два существа, когда они верят, что чувствуют боль?» ответ был таков: в каждом случае, когда существа верят, что они чувствуют боль, это происходит потому, что их мозги находятся в одинаковом физическом состоянии. Если первый ответ представлял собой обычный материалистический тезис, с которым мало кто спорил, второй ответ критики сочли нереалистичным. Ментальные состояние у разных людей не могут быть тождественными, поскольку состояние физического вещества мозга у них неодинаково; у страдающих от зубной боли людей причины

37

боли могут быть разными. Присущий этой теории редукционизм противоречил факту, что люди с

одинаковыми мозгами мыслят по-разному.

Энтузиазм, с каким была встречена теория тождества, довольно скоро сменился разочарованием, и уже к началу 70-х годов у нее было больше критиков, нежели сторонников. Тем не менее данная теория задала ряд базисных параметров для сменивших ее теорий: это перевод нашего рассуждения о ментальном и физическом на языковый уровень и, конечно, физикалистское, материалистическое мировоззрение с его верой в единство науки.

Реакции на теорию тождества были разные. С оригинальной идеей выступил Дональд Дэвидсон, предложив узаконить «аномальный монизм». Ошибка теоретиков тождества, с его точки зрения, состояла в том, что они исходили из существования психофизических законов, дающих возможность в принципе, при достаточно полном знании физических процессов мозга, предсказать конкретное ментальное событие. Таких законов не существует. Они правы, говоря, что каждому нашему ментальному событию соответствуют физические события, но ошибаются в другом: на самом деле ни одно ментальное событие не может быть предсказано, каким бы полным не было наше знание о физических событиях. Их невозможно подвести ни под один естественнонаучный закон. Эту аномальность следует просто принять как данное<sup>1</sup>.

С более радикальных позиций выступили *элиминативисты* — Р. Рорти, П. Фейерабенд, П. Черчленд и ряд других авторов. Элиминативисты считают верной общую стратегию Карнапа на построение единой науки, охватывающей как физическое, так и психическое: Вселенная однородна, непрерывна, каузально связана, действует по единым физическим законам. Однако тактика на трансляцию менталистского языка психологии на язык физики была ошибочной, поскольку

<sup>1</sup> Davidson D. *Mental Events // Experience and Theory*. Ed by L. Foster and J. Swanson. Auherst MA, 1970.

38

первый язык не имеет референтов. На этом же основании ошибочной является и тактика теории тождества: нельзя отождествлять фиктивное с нефиктивным. Рорти предложил радикально изменить тактику: признать, что поскольку используемые в нашей обыденной речи ментальные термины и категории относятся к псевдореалиям, они подлежат элиминации. Прodelать процедуру, аналогичную той, когда из научного знания были элиминированы термины, относящиеся к фиктивным сущностям вроде «флогистона», «дьявольских сил» и т.п. Искать тождество или соответствие псевдореалий процессам мозга — это все равно, что искать референты «дьявольским силам» в анатомии и физиологии организма. Сознание — это только внешне фиксируемые нейрофизиологические процессы, языковое и социальное поведение; здесь нет места для «духа». Элиминация ментальных понятий, по его мнению, освободила бы философию от множества псевдопроблем.

Позицию элиминативистов многие находят логически достаточно последовательной, но контринтуитивной. Признание высказываний с ментальными терминами типа «мне больно», «я думаю», «я верю», «я надеюсь» чистыми фикциями, не имеющими никаких референтов, встречает сопротивление здравого смысла.

Более сильной реакцией на теорию тождества стал *функционализм*. Основная идея функционализма состоит в том, что виды ментальных состояний следует считать не видами физического и вообще не какими-либо *свойствами*, материальными или идеальными, а *функциональными состояниями*. Функционализм существует во множестве версий, у которых есть ряд более или менее общих пунктов согласия. Во-первых, существует «антиредукционистский консенсус»; не требуется ни редукция ментального к физическому, ни объяснение физического через ментальное: сознание рассматривается как чистая *функция*, безотносительно к породившим эту функцию свойствам (поэтому функционалистами могут быть как материалисты, так и идеалисты). Во-вторых, *метафизическая нейтральность*: ментальные и когнитивные свойства иссле-

39

дуются безотносительно к их биологическим, физическим или духовным носителям.

Предполагается, что таким образом можно обойти традиционные опасности как картезианского дуализма, так и редукционистского физикалистского монизма. Ключевыми терминами для функционалистов стали не «ментальное» или «физическое», а «реализация», «имплантация», «воплощение». Данные термины нацелены показать, что ментальное может получать реализацию в разных носителях, не будучи сводимым к ним и не требуя определения их в физических или ментальных терминах. Это значит, феномен, обычно называемый «сознанием», может быть функционально реализован в человеке, в компьютере или в каких-либо других вещах. Подобно тому, как математическая операция с отношением  $2+2=4$  может быть реализована в человеческом мозге, на калькуляторе, на счетах с косточками или присоединением камешков. Это называется «аргументом множественной реализации» (the multiple realisation argument).

Функционалистский подход к сознанию появился в философских дискуссиях в нужное время и в нужном месте. К этому времени он широко и плодотворно использовался в социологии и биологических дисциплинах. Примерно в этот период начинается интенсивное развитие когнитивных наук. Общим термином «когнитивные науки» стали называть междисциплинарные исследования, охватывающие работы по искусственному интеллекту, компьютерной психологии, психолингвистике, нейропсихологии, когнитивной этологии и др. Философский функционализм как нельзя лучше отвечал духу этих новых областей, обрел здесь сторонников среди молодого поколения ученых, став своего рода «ортодоксией».

Хотя в общей форме идея функционалистского подхода к сознанию высказывалась давно, по-настоящему рабочей она стала с внедрением как в психологию, так и в философию компьютерных метафор и обсуждения идеи «Машины Тьюринга», выдвинутой Аланом М. Тьюрингом, британским математиком, известным своими пионерскими работами в компьютерной теории. Одним из первых философов, применившим

40

функционалистско-компьютерный подход к сознанию, был Хилари Патнэм. В 1960 году он опубликовал статью «Сознание и машины», в которой сравнил ментальные состояния с логическими состояниями машины Тьюринга<sup>1</sup>. «Функционализм машины Тьюринга» (Turing machine functionalism), проводящий аналогию между сознанием и компьютерами, открыл широкие перспективы для новых исследований. (Компьютерный подход стал стартовой площадкой для философии сознания Деннета. Более подробно мы будем об этом говорить в следующих главах.)

В силу того что функционализм оказался хорошо работающим инструментом в когнитивных науках, у него оказалась значительно более долгая история, нежели у теории тождества. Тем не менее по прошествии времени возникли вопросы: можно ли понимать «функцию» в отрыве от свойств носителей, в которых она реализуется? Что такое «реализация» и каков ее механизм? Как быть с субъективностью человека — ощущениями, чувствами, эмоциями, которые компьютерная машина не в состоянии имитировать? Возникли подозрения: не является ли замена ментальных свойств функциональными отношениями просто подменой понятий; действительно ли функционализму удалось снять корневую проблему об отношении духовного и телесного, или она просто отодвинута в тень.

Видя изъяны функционализма, некоторые философы пытаются их преодолеть, не покидая его почвы. Одной из таких попыток является концепция «*супервенности*» (от глагола supervening, означающего в английском языке «следовать за чем-то, дополнять прежнее чем-то новым»). Варианты такого рода функционализма предложили, например, Джегвон Ким и Дэвид Чэлмерс. Предполагается, что в таком толковании функционализм позволяет, с одной стороны, сохранить автономный статус ментального и субъективного, а с другой стороны, обойти опасности физикалистского редукционизма. В том же

<sup>1</sup> Патнэм Х. Философия сознания. М., 1999. С. 23—52.

41

русле — в рамках «антиредукционистского консенсуса» функционалистов — происходит возвращение на сцену философии сознания *эмерджентизма*. Не в его классической форме 20—30-х годов, представленной в философии Альфреда Уайтхеда, а с обновленным языком,



проблематикой и техникой. В нем сохраняется главный тезис, что ментальное является новым эмерджентным качеством, возникающим из физического, но не редуцируемым к нему<sup>1</sup>. Идея эмерджентизма является центральной и в книге Джона Сёрля «Открывая сознание заново» (1992)<sup>2</sup>, и у многих других авторов. Однако очень многие критики справедливо замечают, что концепции эмерджентизма создают только иллюзию объяснения, вызывая не меньше вопросов, чем функционализм или теория тождества. Откуда и почему эмерджентизм? У Уайтхеда, широко пользовавшегося этим понятием для обозначения эволюционной новизны, оно было логичным в его системе: под ним было метафизическое основание — идея Бога или творческого процесса Вселенной. Современные эмерджентисты, предпочитающие обходиться без теолого-метафизических понятий, оказываются в затруднении. Каким образом из биологической новации возникает особая бытийственность сознания, способность все «означивать», придавать окружающему миру смыслы и создавать мир артефактов? Мало что проясняет и новейшее словечко «супервенция»<sup>3</sup>.

<sup>1</sup> См.: Bekkermann A., Flohr H. and Kim J. (Eds.) *Emergence and Reduction*. Berlin, 1992.

<sup>2</sup> Сёрль Дж. *Открывая сознание заново*. М., 2002.

<sup>3</sup> Нужно сказать, что идея эмерджентизма весьма популярна среди российских философов, занимающихся проблемой сознания. См.: Дубровский Д.И. Новое открытие сознания? (По поводу книги Джона Сёрля «Открывая сознание заново») // *Вопросы философии*, 2003. № 7; Меркулов И.П. Информационная природа сознания // *Полигносис*, 2000. № 4. Нам кажется, что и Д.И. Дубровский, и И.П. Меркулов, видящие в эмерджентизме «свет в тоннеле», не совсем осознают трудности, заключенные в этой позиции. По сути оба они, как,

42

Суммируя наш очень неполный экскурс полувекового развития англоязычной философии сознания (и упрощая, конечно, картину), мы видим ее историю как умножение позиций в толковании феномена сознания. Это подтверждается количеством фигурирующих в ней «измов» — «материализм», «физикализм», «функционализм», «элиминативизм», «функционализм машины Тьюринга», «информационная теория», «семиотический материализм», «дуализм», «дуализм свойств», «натуралистический дуализм», «коннективизм», «двухаспектная теория», «аномальный монизм», «эмерджентизм», «супервенциализм», «эпифеноменализм», «когнитивизм» и др. В. Ленин когда-то пренебрежительно оценил множественность позиций в философии опыта (и их взаимную критику), как свару «желтых и зеленых чертей». Это, конечно, поверхностный взгляд на характер философских поисков. На самом деле за плюрализмом скрывается и особый статус загадки сознания, огромные (и особые) трудности ее исследования, и опробующий поиск — в конкурентной борьбе — возможных путей ее решения.

В плюрализации позиций в философии сознания тем не менее есть ряд определяющих векторов поисков. Прежде всего, вектор, сложившийся вокруг альтернативы социологизм—биологицизм. Отсутствие в науке твердо установленных фактов относительно генезиса сознания делает ее особо острой. Можно ли феномен осознающего себя интеллекта сводить только к содержательной, когнитивной стороне и объяснять его лингвистическими и социальными факторами? Если это так, то как

впрочем, и зарубежные биологицисты Дж. Сёрль, А. Кларк, П. Черчленд, Дж. Ким и др., сделали выбор в пользу позиции, согласно которой сознание является биологической инновацией. Поскольку, как говорит Меркулов, «сугубо культурная эволюция человека как интеллектуального вида практически невероятна» (с. 27). Однако при этом остается открытым вопрос о том, каким образом у биологической материи появляется означивающая способность. Теории информации, коннективистские или иные концепции мало проясняют картину апелляцией к информационным системам природы.

43

именно эти факторы осуществляют свою магию, в результате чего биологическая материя мозга превращается в сознание? Или этот феномен все же порожден биологически базисными процессами природы (или ее информационными структурами), в примитивном виде присутствуя в психике животных, а язык и социальное обучение только оформляет и шлифует его? Если верно



последнее, возникает другое недоумение: что представляют собой эти базисные процессы и почему только у человека они превратились в осознающий себя интеллект? Разумеется, никто из авторов не отрицает, что сознание творится и в природной и культурно-языковой среде. Спор идет по поводу того, какую исследовательскую стратегию считать оптимальной в генезисе и какие факторы принимать за приоритетные — биологические или социологические. Одни считают, что истоки сознания следует искать в биологических информационных структурах и животной психике, другие отводят доминантную роль искусственно созданным семиотическим системам<sup>1</sup>.

<sup>1</sup> В российской философской литературе 60—70-х годов тоже сложились аналогичные векторы. Примером радикального социологизма может служить статья Э.В. Ильенкова «Идеальное» в «Философской энциклопедии» (М., 1962). В ней говорится, что мышление «есть не индивидуально-психологический, тем более не физиологический факт, а факт общественно исторический, продукт и форма духовного производства» (с. 219). В отличие от зарубежных социологов, согласных с тезисом Маркса, что материальной действительностью мысли является язык, и отождествляющих сознание с социолингвистическим поведением, Ильенков не считает язык определяющим фактором: главное в мышлении — это «особая функция общественно-трудовой деятельности» (с. 221). Лингвистически ориентированные философы, с его точки зрения, совершают такую же натуралистическую ошибку, как и те, кто отождествляет мышление с биологией и физиологией (там же). Согласно Д.И. Дубровскому, защищающему информационный подход к сознанию, сознание возникает как эмерджентное качество биологической материи, превращаясь в обществе и культуре в самоорганизующуюся систему. (Дубровский Д.И. Информация, сознание, мозг. М., 1980.)

44

От принятия позиции в отношении генезиса сознания во многом зависит отношение различных авторов к объяснительным возможностям искусственного интеллекта. (Здесь тоже существует свой раскол: одни видят в компьютерах долгожданное средство прояснения тайны сознания, другие оценивают вторжение компьютерных метафор в философию как бедствие.) Огрубляя ситуацию, можно сказать, что в современной философии сознания фигурируют несколько стратегий, которые можно было бы назвать социолингвистической, биолого-социологической и социобиологической. (Последняя стратегия характерна для Деннета. Как мы покажем дальше, все проблемы он решает, имея в виду двойную перспективу искусственного и естественного.)

Другой важный вектор сложился вокруг старой дилеммы монизма—дуализма. Образно говоря, спорящие разделились на два лагеря: одни считают, что следует окончательно изгнать «Дух из Машины», другие, напротив, убеждены, что настала пора снова загнать «Дух в Машину». Для философов первого лагеря, изгоняющих «Дух из Машины», как и во времена неопозитивизма, эталоном объяснения реальности является наиболее отстраненная от субъективного и ценностного взгляда на мир наука— физика; все, что считается ментальным, поддается объективному (как в науке) объяснению. Философы из второго лагеря выражают сомнение относительно эффективности физикалистских (в том числе компьютерных) стратегий; при их применении к сознанию выпадает очень важный факт — наше уникально человеческое отношение к миру.

Примечательная черта нынешнего расклада сил состоит в том, что новое дыхание получили дуалистические, или, точнее, неодуалистические концепции. Все громче звучат голоса тех, кто настаивает на реабилитации феноменального мира внутреннего опыта и, соответственно, на восстановлении в правах самоотчета субъекта о его собственном Я. Говорят о необходимости восстановления в правах самости как первоисточника творчества и значений (и культуры), и субъективной реально-

45

сти (понимаемой по-разному). Однако в этом случае остается неясным, как можно состыковать онтологию субъективности с наукой. Одним словом, эпистемологический вопрос о понимании принципов объективности и субъективности остается по-прежнему актуальным. (Радикалистские идеи Деннета, подлив масла в огонь дискуссий, спровоцировали многих философов защищать дуализм с открытым забралом.)<sup>1</sup>

Конечно, физикалистский материализм никогда не имел недостатка в оппонентах. В 60—70-е годы, в самый разгар его наступления на «бастион сознания», с резкой критикой монистического

физикализма и защитой дуализма выступили Карл Поппер и Джон Экклз<sup>2</sup>. Однако при царившем тогда оптимизме относительно возможности объяснения сознания, обходясь без ментального, их сочли консерваторами. В последние десятилетия голоса в пользу реабилитации или «нового открытия сознания» высказывают такие авторитетные философы, как Джон Сёрль, Колин МакГинн, Томас Нагель, Джон Фостер, Майкл Локвуд, Нэд Блок, Сол Крипке, Роберт Пенроуз, Овен Фланаган, Дэвид Чэлмерс и др. Конечно, речь идет не о реабилитации чреватого идеализмом субстанциалистского декартовского дуализма. Скорее — о необходимости поиска нового объяснения сознания и нового понимания объективизма, могущего включить в себя

<sup>1</sup> Умножение загадок феномена сознания и увеличение трудностей его объяснения стимулировало появление новых форм философской теологии, предлагающих альтернативные материализму объяснения. В их числе теистически-натуралистические формы. В рамках «теистического натурализма», например, утверждается, что объяснение феномена сознания (как и материи) требует принятия посылки о «Разумном Дизайне» (Intelligent Design). Под «дизайном» понимается не традиционный моральный Бог, а Бог как процессуальная работа Вселенной или функционирование структур самоорганизующихся систем. Соответственно новое дыхание получили теистические варианты телеологии.

<sup>2</sup> Popper K., Eccles J. *The Self and its Brain*. In *Defence of Interactionism*. B., N.Y., L., 1977.

46

субъективность или «точку зрения от первого лица». Дэвид Чэлмерс назвал это движение «натуралистическим дуализмом». В стойкости «бастиона сознания» перед атаками материалистов и функционалистов многим видится свидетельство того, что скроенная по меркам физических наук объективистская методология является ограниченной и что следует продолжить поиски более адекватных ее моделей. О новом драматическом витке напряженности и об обострении конкуренции концепций сознания можно судить по публикациям последнего времени<sup>1</sup>.

Оценивая полувековые атаки на «бастион сознания», было бы неверно утверждать, что эти атаки захлебнулись или были безрезультативными. Если оценивать прогресс в философии сознания с точки зрения прояснения проблем, которые со времен Декарта виделись смутно, а обсуждались спекулятивно, то он налицо. Какие-то подходы признаны несостоятельными, а какие-то перспективными. Расширилась проблематика и более

<sup>1</sup> Перечислим только некоторые, получившие широкую читательскую реакцию: Penrose R. *The Emperor's New Mind: Concerning Computers, Minds and the Laws of Physics*. Oxford, 1985; Nagel Th. *The View from Nowhere*. N.Y., Oxford, 1986; Block N. *The Computer Model of the Mind*. 1990; Dennett D. *Consciousness Explained*. Boston, 1991; Searle J.R. *The Rediscovery of Consciousness*. Cambridge MA, 1992; McGinn C. *The Problem of Consciousness*. Oxford, 1990; Lockwood M. *Mind, Brain and the Quantum: the Compound I*. Oxford, 1991; Flanagan O. *Consciousness Reconsidered*. Cambridge MA, 1992; Kim J. *Supervenience and Mind*. Cambridge, 1993; Churchland P.M. *The Engine of Reason. The Seat of the Soul*. Cambridge MA, 1995; Dretske F. *Naturalizing the Mind*. Cambridge MA, 1995; Chalmers D. *The Conscious Mind. In Search of a Fundamental Theory*. Oxford, 1996; Kirk R. *Raw Feeling. A Philosophical Account of the Essence of Consciousness*. Oxford, 1996; Lycan W.G. *Consciousness and Experience*. Cambridge MA, 1996; Clark A. *Being There: Putting Brain, Body and World Together Again*. Cambridge MA, 1997; Searle J. *Mystery of Consciousness*. N.Y., 1998; Chomsky N. *New Horizons in the Study of Language and Mind*. Cambridge MA, 2000; Crane T. *Elements of Mind*. Oxford, 2001.

47

разнообразными стали когнитивные инструменты. Конечно, философия сознания столкнулась с новыми сложными проблемами. Такое происходит со всякой развивающейся областью знания; за победами открываются новые загадки, а вместе с ними возникает потребность в новых объяснительных стратегиях. Тем не менее все это следует, как нам кажется, рассматривать как признаки прогрессирующего движения в философии.

В завершение нашей краткой зарисовки истории вопроса — драматической осады «бастиона сознания» — мы выражаем надежду, что читатель более подготовленно воспримет дальнейший разговор о концепции сознания и самости Деннета. Нам хотелось, чтобы он почувствовал, что она «варилась» и «созревала» в поле этой драматической напряженности. Не будь этой напряженности

(и сильных конкурентов в его окружении), его философия либо не состоялась бы, либо состоялась в урезанном виде.

Всякая претендующая на новизну и охват широкого круга проблем теория, как правило, имеет критическую и конструктивную стороны. В первой части книги мы проанализируем обе эти стороны. В главе II рассматривается деконструктивная часть философии Деннета — его опровержение феноменологии (или феноменализма), как необходимого условия преодоления дуализма. В главе III рассказывается о его конструктивистском проекте — Модели Множественных Набросков — и его версии функционализма. В главе IV разбирается концепция виртуальности сознания в свете эволюционной теории Дарвина и «Универсальной машины Тьюринга». Главы V, VII и VIII посвящены проблематике человека: в них рассказано о понимании Деннетом самости, личности, свободы воли, а также о его взглядах на отношение механистического и целеполагающего объяснений человека. Глава VI целиком отведена рассмотрению теории интенциональности, играющей основополагающую роль во всей его философии. Вторая часть книги посвящена полемике Деннета с его оппонентами и коллегами. Заключение содержит ряд наших выводов и наблюдений, касающихся как концепции Деннета, так и современной англоязычной философии сознания.

48

## II. Деконструкция опыта и сознания

Защищать квалиа от гетерофеноменологических поползновений можно только ценой утверждения, что они совершенно недостижимы для науки.

*Дэниел Деннет*

Забудьте культуру, забудьте язык. Тайна начинается с самого низшего организма, который, если вы ткнете, например, в него булавкой, не просто реагирует, но действительно *чувствует* что-то.

*Майкл Локвуд*

### 1. Иллюзии «Картезианского Театра»

Главный стратегический замысел Деннета состоит в том, чтобы преодолеть идущий от Декарта и Brentano раскол мира на *res extensa* и *res cogitans*, на сферу физического и сферу интенционального и обосновать непрерывность мира и единство знания. Камень преткновения, о который спотыкались все предпринимавшиеся до этого попытки, — феномен сознания. Не убрав его с дороги, не объяснив сознание в рамках непрерывности мира, нельзя говорить о единстве знания. Неудачи прежних попыток он объясняет тем, что новаторы вели поиск вариантов решений в оппозиции физикализма и ментализма и использовали неработающие редуccionистские модели. В их распоряжении еще не было инструментов, которые сегодня предлагают новые философские теории, когнитивные науки,

49

науки об искусственном интеллекте и др. Деннет убежден в возможности третьего — *нейтрального* — пути между жестким элиминативистским физикализмом, с одной стороны, и ментализмом, с другой, и эта убежденность у него основывается на применении к исследованию сознания этих новых инструментов.

Основным объектом его критики является представление о сознании как своего рода *homunculi*, спрятанном в голове человечке (или «Внутреннем Оке», «Главном редакторе», «Боссе»), который изнутри наблюдает поток ощущений и представлений, сортирует их и отдает команды поведению. Образ сознания, который лепился в традиции Декарта, Brentano и Гуссерля, Деннет именуется «Картезианским Театром»<sup>1</sup>. В нем есть и сценическое действие — поток ощущений, восприятий, мыслей, и зритель — *homunculi*. Крайняя форма гомункуляризма, воплощенная в теле декартовская духовная субстанция.

Такой образ сознания и самости, убежден Деннет, является совершенно ложным; он не оправдан философски и пришел в противоречие с имеющимися научными фактами об интеллектуальной деятельности человека. Исходный ложный постулат, на котором строится идея Картезианского

Театра, — дуализм, отнесение сознательных процессов к *особым ментальным, субъективным, внутренним* состояниям. Данный постулат предполагает некую «добавку» к физической Вселенной. Верный духу позитивизма и идеалу физикализма, Деннет убежден, что мир един, а не раздвоен, в нем действует единая система каузальных связей, соответственно сознание и самость должны объясняться в рамках единого, объективного знания.

Сила дуализма состоит в том, что он является естественной и привычной позицией человека в отношении своего

<sup>1</sup> Метафору «театра» Деннет заимствовал у Юма, когда тот сравнивал сознание с театром, в котором происходит последовательная смена восприятий.

50

интеллектуального поведения. Дуализм составляет основу повседневного самовосприятия, здравого смысла и всего того, что называют «фолк-психологией». Слыша шум дождя, человек верит, что внутри него шум представлен в виде некоего ментального, но вполне реального качества. Ощущая, переживая, обращаясь к образам памяти или намереваясь что-то предпринять, он верит, что его ощущения, переживания, образы и намерения имеют отношение к находящейся в центре его мира *директивной самости*. Одним словом, он верит в существование внутри него некоего *медиума*, репрезентирующего и внешне физические вещи, и внутренние ощущения тела, и объекты воображения. Таким медиумом не может быть серое вещество мозга. Рассматривая мозг в лаборатории, он видит только физическое тело; следовательно, сознание должно быть какой-то качественно иной добавкой к этой материальной субстанции.

Несмотря на естественность такого представления человека о самом себе, дуализм, утверждает Деннет, должен быть отброшен и даже «догматически отвергнут». В «Сознании объясненном» он пишет: «...в этой книге я принял внешне догматическое правило, что дуализма следует избегать любой ценой»<sup>1</sup>. Исследователям он дает совет — никогда не поддаваться под очарование песни, которую поет Сирена дуализма. Это не значит, говорит он, что у меня в кармане стопроцентное опровержение дуализма. «Я совершенно уверен, что никакое такое опровержение невозможно, но это слабое утешение для дуализма..»<sup>2</sup>. Дуализм должен быть лишен права быть одной из объяснительных возможностей в силу фатальных пороков, делающих его «безнадежным». (Один из разделов книги «Сознание объясненное» так и называется «Почему дуализм безнадежен?») Во всех своих работах, касаются ли они проблем свободы и детерминизма, интенциональности и

<sup>1</sup> Dennett D. *Consciousness Explained*. Boston, 1991. P. 37. <sup>2</sup> Dennett D. *Living on the Edge*. «Consciousness Explained». Symposium // *Inquiry*. V. 36. № 1/2. 1993. P. 140.

51

рациональности, естественного и искусственного интеллекта, механистического и целеполагающего объяснения, он не устает приводить все новые и новые аргументы, долженствующие показать, почему не следует присушиваться к «Сирене дуализма». Мы будем разбирать их в следующих главах, сейчас же по пунктам обозначим главные.

— Для научно ориентированного человека дуалистическая версия отношения духовного и телесного неприемлема потому, что она противоречит естественнонаучным фактам, вера в которые имеет прочные основания. В случае дуализма сознание оказывается вне досягаемости объективного исследования: нельзя исследовать то, что не поддается интерсубъективному анализу или обозначается с помощью сакраментальной фразы: «А здесь имеет место чудо». В частности, дуализм не сообразуется с новыми данными когнитивных наук, не говоря уже о его негативном влиянии на всю исследовательскую культуру своим расколом знания на «науки о духе» и «науки о природе».

— Дуализм неприемлем по философским соображениям. Существенный изъян позиций дуалистов, или гомункуляристов, состоит в том, что гомункулус нуждается в объяснении — откуда он взялся? Если исходить из религиозного объяснения, как это было с Декартом, тогда вопрос об источнике получает видимость решения; если же не принимать его, возникает *угроза бесконечного регресса*. Для понимания сознательных процессов требуется «Театр» с его «аудиторией», а для понимания «Наблюдающего режиссера» требуется постулировать другой

«внутренний театр», другого «режиссера» и другую «внутреннюю аудиторию». И так до бесконечности. Некоторые философы полагают, что избежать этой угрозы невозможно. Деннет с ними не согласен. Положить конец регрессу можно, если отказаться от сущностной идеологии и принять организмическую (эволюционную) тактику. Способ разрешения парадокса регресса такой же, какой он и при определении сути млекопитания. Например, каждое млекопитающее имеет мать, и это подразумевает бесконечную генеалогию млекопитающих (что невозмож-

52

но). На самом деле нам следует искать не сущность млекопитания, что позволило бы нам в принципе идентифицировать «Первого Млекопитающего», но терпимо относиться к конечному регрессу, связывающему млекопитающих с их не-млекопитающими предками, с последовательностями, которые могут быть только частичной произвольностью<sup>1</sup>. Самое главное в объяснении сознания без регресса — отказаться от заглядывания вовнутрь и сосредоточить внимание только на том, что поддается описанию «с позиции третьего лица», то есть intersубъективно.

— Приписывая ментальной самости статус центра, вокруг которого разворачиваются сценические действия сознания, дуализм наделяет самость *самоактивностью, или самотворчеством*. В самоактивности самости усматривается смыслообразующий источник всех артефактов и всего многообразия культуры. Она наделяется статусом «Центрального означивателя» (Central Meaner), которого никто не означивал, но который сам способен придавать всему значения (Non-meant Meaner). «Центральный означиватель» играет такую же роль, какую в космологии Аристотеля играл «неподвижный двигатель»; находясь в покое, он задает импульсы всем видам движения.

— Эпистемологическая порочность дуализма обусловлена тем, что он апеллирует к *интроспективизму*, идее, что содержание сознания открыто только внутреннему наблюдению агента в ходе его «прямого» феноменального опыта и непосредственно им удостоверяется. Между тем и в философских и в научных исследованиях накоплено большое количество достоверных фактов, говорящих об обманчивости интроспекции и недостоверности «непосредственно данных опыта».

— Одной из составляющих дуализма является *интенциональный реализм*. Интенциональность полагается внутренним свойством ментального, проявляющегося в направленно-

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 144.

53

сти сознательных актов на что-то, выступающего для нас объектом хотения, желания, целеполагания. Идея внутренней интенциональности, как и внутренней интроспекции, обманчива. На самом деле интенциональность есть внешняя установка интерпретации, исходя из которой мы оцениваем объекты, приписывая им интенциональные свойства.

— Дуализм исходит из идеи *психофизической интеракции*, апеллирующей к неведомой науке психофизической каузальности. Поскольку интеракция не поддается никакой объективной фиксации, на деле это есть апелляция к чуду. Данное стандартное возражение интеракционизму было знакомо еще самому Декарту, но ни он, ни последующие дуалисты никогда не сформулировали убедительного ответа. Особенно загадочной выглядит директивная способность нематериального сознания отдавать приказы материальному мозгу и телу. Согласно дуалистической гипотезе, эта способность не является физической: ни световыми излучениями, ни звуковыми волнами, ни потоком субатомных частиц, с ней не ассоциируется никакая физическая энергия, и все же она каузально воздействует на тело. Фундаментальный принцип физики состоит в том, что любое изменение в траектории какой-либо физической сущности есть ускорение (или замедление) и затрата энергии. Но откуда берется эта энергия у сознания? Чем она может быть? Интеракционизм не в состоянии ответить на этот вопрос. Любое предположение о каузальном влиянии нефизического сознания на физические изменения тела противоречит фундаментальному принципу сохранения энергии. «Именно принцип сохранения энергии объясняет физическую невозможность существования «машин с вечным двигателем», а дуализм нарушает этот самый принцип. Конфронтация между вполне стандартной физикой и дуализмом была предметом



бесконечных дебатов после Декарта, и повсеместно принято считать, что это — неустранимый и фатальный порок дуализма»<sup>1</sup>.

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 33—35.

54

Одним словом, Деннет убежден, что даже этот небольшой перечень «фатальных пороков» дуализма делает его неконкурентоспособным в сравнении с физикалистским монизмом, утверждающим, что физический мир представляет собой *закрытую систему*, в которой действуют единые каузальные зависимости.

## 2. «Картезианский материализм»

Поставив перед собой цель развенчать образ сознания в виде «Картезианского Театра», смысловые действия которого зависят от «исполнительного босса» — самости, Деннет имеет в виду не столько позиции философов прошлого, сколько позиции современных его коллег. Многие из них и сегодня вовсе не склонны отказываться от идеи реальности и центризма самости, хотя и признают ее удивительную неуловимость для анализа. (Разумеется, они не отрицают роль языка и культуры в формировании самосознания. Однако самоактивная самость все же полагается ее изначальным условием.) Чаще всего защита гомункуляризма связана с сомнениями относительно возможностей панобъективизма в объяснении феноменальных и субъективных качеств опыта. Сомнения выражаются в принятии тезиса, что сознательные процессы являются особыми внутренними свойствами индивидов. Объяснение ментальной функции в «экстерналистских» терминах (процессов мозга, лингвистического поведения и др.), считают они, всегда неполно, поскольку в нем пропадает аспект интеллектуальной жизни, который не является физическим или лингвистическим, выражая индивидуальное видение мира «с позиции первого лица». Поэтому объективное описание опыта требует обращения к внутреннему опыту. (Такую позицию называют «интернализмом».)

Деннет — убежденный «экстерналист». Скептиков, сомневающихся в потенциях объективистских методов, называет

55

«дуалистами» или «полудуалистами», «скрытыми дуалистами». Одни известные авторы, говорит он, защищают дуализм открыто: Артур Кестлер, например, провокативно — в пику Дж. Райлу — назвал свою книгу «Дух в машине» (1967). Многие из этих авторов по своим общим взглядам являются натуралистами и поклонниками науки. Когда же речь заходит о сознании, они сдают свои позиции. Карл Поппер и Джон Экклз, озаглавив книгу «Самость и ее мозг. Аргумент в пользу интеракционизма» (1977), сделали мозг «принадлежностью» самоактивной самости, тем самым отступили от натурализма в сторону дуализма. Одним из активных защитников дуализма является Джон Фостер, полагающий, что говорить о несостоятельности дуализма преждевременно, пока наука не объяснила досконально, какие процессы происходят в мозгу при производстве сознания. Т. Нагеля и Дж. Сёрля — активных защитников субъективности — Деннет называет «скрытыми дуалистами». Сравнительно недавно попытки спасти дуализм предпринимались с другой стороны — со стороны истолкователей физических явлений, прежде всего квантовой физики. Приводились искусные технические доводы, основанные на софистичном прочтении соответствующих физических концепций и толковании активной роли наблюдателя. В свое время они широко обсуждались, однако не обрели большого количества сторонников, поскольку, как считает Деннет, в философском отношении они оказались слабыми.

Примечательная сторона деннетовской атаки на дуализм состоит в том, что объектами критики являются не только менталистский вариант дуализма, но и его материалистический двойник, который он называет «*Картезианским материализмом*». Некоторые его сторонники, например, Майкл Локвуд, уповают на нахождение в мозгу скрытого тайника, где прячется сознание. Деннет считает, что в той или иной форме «картезианские материалисты» воспроизводят идею «шишковидной железы» Декарта, чаще всего в виде «главного нейрона» — физического аналога ментальной самости, стягивающе-

56

го к себе чувственную информацию, редактирующего ее и дающего команды поведению. Этому нейрону, или «центру», приписывается великое таинство — интеракция сознательных процессов с нейрофизиологическими процессами мозга. Хотя для «картезианских материалистов» остается загадкой, где и как в мозгу может происходить встреча мысли и физического вещества, они не отступают от своей стратегии найти точную локализацию этой встречи. По Деннету, эта стратегия заранее обречена на провал: объяснение сознания в терминах физических процессов мозга всегда существенно неполно, поскольку упускается главное — содержательный аспект сознания, процесс движения мысли. Примечательна его оценка такого рода материализма: «Ученые и философы могут прийти к той или иной форме консенсуса в пользу материализма, но освободиться от старых дуалистических образов значительно труднее, нежели это полагают современные материалисты... Это требует весьма пугающей корректировки наших привычных способов мышления...»<sup>1</sup>.

Иначе говоря, Деннет не приемлет ни образ сознания в виде *res cogitans*, ни образ сознания в виде *res extensa*. Однако игнорирование сознания для него тоже неприемлемо. Последнее характерно не только для философов-элиминативистов, полагающих, что за понятием сознания скрывается фикция, но и когнитивных ученых; занимаясь конкретными вещами, они стараются избегать вводить в свои схемы то, что философы именуют «сознанием». В выпадении сознания из их исследовательских программ повинны и сами философы, точнее, их равнодушие к открытиям, сделанным в этих науках, убеждение, что сознание — это не вотчина ученых. Неудивительно, говорит Деннет, что когнитивным ученым «сознание представляется последним бастионом оккультных свойств, эпифеноменом, неизмеримыми субъективными состояниями, одним словом, той частицей сознания, которую лучше оставить философам, принимающим ее с распростертыми объятиями.

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 37.

57

Пусть, мол, они дурачат себя, пытаясь загнать ускользающее нечто «феноменологии» в респектабельную теорию»<sup>1</sup>.

Имея в виду разделение труда, говорит Деннет, в таком отношении когнитивных ученых к философам не было бы ничего страшного, если бы они со своей стороны предлагали приемлемое объяснение взаимодействия между используемыми ими моделями и тем, чем, по их предположению, должно быть сознание. Однако такое объяснение, как правило, отсутствует. Для многих нынешних ученых-когнитивистов вопрос о сознании — пустой вопрос. Хотя на самом деле в науках до сих пор не совершено открытий, которые могли бы если не снять, то хотя бы пролить существенный свет на какие-то философские проблемы. Натываясь на вопросы, ответить на которые когнитивисты, исходя из собственных ресурсов, не в состоянии, они часто попадают в ловушки устаревших философских клише. Деннет убежден, что научные теории об интеллектуальной деятельности никогда не перестанут нуждаться в философии. Даже в предполагаемом Золотом веке нейрофизиологии, когда досконально будет изучено каждое нервное волокно, студенты не перестанут задавать вопрос: «А что такое сознание?» Взаимное непонимание и отчуждение философов и когнитивистов отрицательно сказывается на деятельности и тех и других. Поэтому одну из своих задач он видит в том, чтобы, используя философские инструменты, «заполнить эту пропасть и обрисовать теорию сознания, которая может быть продолжением когнитивистских теорий, и оказать им помощь в объединении нынешних когнитивистских теорий восприятия, решения проблем и использования языка»<sup>2</sup>. Предполагается, что такая теория не будет вступать в конкуренцию с нейрофизиологическими и психологическими теориями. Ее роль другая — служить концептуальной основой когнитивных концепций и источником фундаментальных ме-

<sup>1</sup> Dennett D. *Brainstorms. Philosophical Essays on Mind and Psychology*. Montgomery, 1986. P. 149.

<sup>2</sup> *Ibidem*.

58

тафизических (и эвристических) гипотез. И в этом отношении она будет не подменять их, а

дополнять.

О понимании Деннетом взаимоотношения когнитивизма и философии мы будем говорить в следующих главах. А сейчас обратимся к важнейшей части его деконструктивистского проекта — разрушению «фенома» или «дисквалификации квалиа».

### 3. Миражи «феноменологического сада»

Решающую роль в наступлении Деннета на дуализм играет опровержение феноменализма (или феноменологии); не доказав ложности представления об особой реальности феноменальных качеств и их прямого осознания во внутреннем опыте (внутренней репрезентации), нельзя опровергнуть иллюзии Картезианского Театра. Продолжив линию Л. Витгенштейна, Н. Гудмена, У. Селларса, У. Куайна и других философов-аналитиков, разоблачавших «миф данного», он подвергает дальнейшей деконструкции традиционное представление об опыте.

В самом термине «феноменология» Деннет не видит ничего предосудительного. Еще в XVI веке он использовался учеными для нейтральной и дотеоретической каталогизации всего того, что они обнаруживали в опытах, например, в физических опытах с магнетизмом. Нынешние когнитивные ученые тоже занимаются феноменологией в том смысле, что стремятся как можно более объективно описать все то, что они находят в нем. Совершенно иной смысл этому термину придала философская школа феноменологии, зачинателем которой был Гуссерль; термин стал применяться для *дотеоретической* дескрипции того, что мы находим в субъективном потоке сознания при условии выведения внешнего мира «за скобки». В этой традиции господствует точка зрения, что сознание не поддается объективному исследованию «со стороны» и что для его постижения требуется прямое обращение субъекта к его внутреннему миру.

59

В 1979 году Деннет опубликовал статью «Об отсутствии феноменологии»<sup>1</sup>, в которой писал, что наши отчеты о том, что лично нам кажется собственным ментальным миром, настолько отличны от того, чем в действительности является этот мир, что для его описания привычный феноменальный язык не годится. В субъективном мире человека могут присутствовать множество образов, событий, звуков, запахов и других чувств, в отношении которых он искренне верит, что они существуют в его потоке сознания в виде «флоры и фауны феноменологического сада» или «квалиа». В статье он постарался показать, что на самом деле в природе нет такой вещи, как квалиа. Не существует внутренних феноменов, которые кажутся наличествующими в мире нашего сознания, нет и смыслообразующей активности нашего Я. При пристальном исследовании «окажется, что то, что проплывает в потоке сознания и известно нам как боль, ароматы, мечты, ментальные образы, вспышки гнева и вожделения — стандартные обитатели фенома, — вовсе не являются таковыми, какими они нам кажутся. В действительности они столь отличны от кажущегося, что для них следует найти какие-то новые слова»<sup>2</sup>. В «Сознании объясненном» он вспоминает, что высказанный в статье совет — выбросить *phenom* из объяснительных схем — многие коллеги тогда восприняли как противоречащий здравому смыслу и верх нелепости. Из этого он сделал вывод, что требуется дополнительная работа по развенчанию иллюзий кажимости привычного — феноменального — самовосприятия внутреннего мира.

По поводу онтологии феноменального современных философов условно можно разделить на два лагеря: ее противников и защитников. В число первых входит Витгенштейн. Витгенштейн поставил под сомнение достоверность приватного опыта и отрицал возможность корреляции значений психологических предикатов (боль, красное) с чем-то, находящим-

<sup>1</sup> Dennett D. On the Absence of Phenomenology // *Body, Mind and Method*. Ed. by D. Gustafson and B. Tapscott. Dordrecht, 1979.

<sup>2</sup> Dennett D. *Consciousness Explained*. P. 45.

60

ся в субъективности. Подчеркнув логическую приоритетность публичного языка в сравнении с приватным, он отверг представление о языке как носителе независимых от языка ментальных качеств: приватный опыт не может выступать критерием правильности данных нам в нем качеств.

Он предложил искать референты психологических предикатов в специфике контекста их применения: в обстоятельствах, в которых то или иное поведение оправдывает их приписывание субъектам. Идеи Витгенштейна составили костяк социолингвистической парадигмы. Бихевиористски ориентированные философы нашли перспективным этот подход. Это относится и к Деннету, хотя, как мы покажем далее, он внес в него существенные натуралистические коррективы.

Сомнения, посеянные Витгенштейном относительно достоверности приватного опыта, его подчеркивание приоритетности публичного языка и атаки постпозитивистов на идею фундаментальности данных опыта получили поддержку с другой стороны — со стороны когнитивной науки, в том числе и когнитивной психологии. Следует сказать, что термин «когнитивизм», получивший в них широкое хождение, не имеет однозначного определения. (Что в немалой степени связано с отсутствием согласия по поводу основополагающих принципов работы в этой области и ключевого понятия — «сознание».) В когнитивной психологии, возникшей как реакция на прямолинейный бихевиоризм, под «когнитивизмом» часто понимается признание первостепенной роли содержательной, информационной компоненты в психологической детерминации поведения<sup>1</sup>. В когнитивных науках, сердцевину которых составляет теория искусственного интеллекта, «когнитивизм» означает фиксацию в сознательном опыте тех сторон, которые в принципе поддаются компьютеризации и объективной репрезентации. Мыслительный процесс рассматривается как вычислительный процесс, бессознательный механизм пе-

<sup>1</sup> См.: Величковский Б.М. Современная когнитивная психология. М, 1983.

61

реработки информации, который не сводится ни к нейрофизиологии мозга, ни к феноменологической стороне опыта. Последняя в них чаще всего игнорируется.

И все же многие философы находят контринтуитивную позицию отрицания феноменологии. Витгенштейн, говорят они, не прояснил, а обошел многие сложные вопросы. Личные переживания удостоверяются приватно, а не публично. Некоторые философы делают еще более решительные заявления. Они утверждают, что феноменальные качества опыта, или квалиа, существуют и что настала пора спасти их из-под лавины языковых, информационных и когнитивистских интерпретаций. (Термин «квалиа» — от лат. *qualia* (*quale*) — «качества», «свойства», «предмет обсуждения» — в силу его краткости стал в современной литературе наиболее употребительным). Состояние сознания всегда предполагает, чтобы что-то выглядело как что-то, представало в виде определенного качества. Без этой посылки всех нас можно считать зомби: если у нас нет восприятия чего-либо в виде качественных состояний, у нас вообще нет состояния сознания. В свою очередь, противниками онтологизации квалиа был высказан такой контраргумент: экспликация одного термина в терминах другого — это движение по кругу, ибо она не проясняет, что такое быть осознающим бытием или субъектом сознательного опыта. Одним словом, в контроверзе по поводу феноменальных квалиа новый смысл приобрел тезис Беркли «*esse est percipi*»: для многих он стал означать, что быть в сознании — значит иметь квалиа.

В 1983 году Дж. Левайн опубликовал получившую широкий резонанс статью «Материализм и квалиа: пропасть в объяснении»<sup>1</sup>, в которой привел ряд доводов о наличии пропасти между когнитивной и феноменальной сторонами опыта сознания. Он утверждал, что даже если нам удастся когда-либо объяснить функционирование когнитивной (презентативной)

<sup>1</sup> Levine J. Materialism and Qualia: The Explanatory Gap // *Pacific Philosophical Quarterly*. V. 64. P. 354—361.

62

системы сознания, мы не приблизимся к объяснению того, какой нейробиологический механизм во взаимодействиях с окружающими условиями в итоге генерирует квалиа. Нэд Блок (Массачусетский институт технологии) считает, что в отношении феноменальных сторон (способов, какими вещи выглядят для нас, пахнут, звучат) и когнитивных (рефлексивных, способных к репрезентации) сторон сознательного опыта существует так много загадок, что говорить об элиминации феноменального еще рано. «Ни одно из имеющихся сегодня в наличии

нейрофизиологических или компутационных понятий не в состоянии объяснить, что же это такое — быть в феноменально осознанном состоянии, то есть иметь боль или видеть что-либо красное. ...Это хорошо известная пропасть в объяснении»<sup>1</sup>. Идея «пропасти» присутствует и в книге Дэвида Чэлмерса «Осознаваемое сознание» (1996). Чэлмерс выделил в проблеме сознания «легкий» и «трудный» аспекты. Когда что-либо, например, палка, выглядит определенным образом (выступает как квалиа), его репрезентация в нашей когнитивной экономике может быть двойкой. Первая роль — когнитивная или содержательная — логически связана с другими репрезентациями, с верованиями, с памятью, с действием и др. Для когнитивной презентации не имеет значения, например, осознаю ли я палку или нет. Вторая

<sup>1</sup> Block N. Review on Daniel Dennett's «Consciousness Explained» // Journal of Philosophy. V. XC. № 4. April 1993. P. 182. Блок пишет: «Предположим, например, что Фрэнсис Крик и Кристоф Кох правы, что короткие, в пределах 35—75 герц, колебания нервных волокон являются существенными для феноменального сознания. Никому еще не удалось обрисовать объяснительную связь между этими нервными колебаниями и феноменальным сознанием. Почему 35-75 герц, а не 5 герц или 100 герц? Далее, как один тип колебания, скажем в 40 герц, может объяснить, почему получаемое в результате осознаваемое состояние ощущается одним способом, нежели другим, или вообще не чувствуется? На данный момент ничего, что нам известно из нейрофизиологии или компутационной психологии, не дает ключ к ответу на эти вопросы» (ibidem).

63

роль состоит в осознанности палки как палки, в ее явлении мне качественно определенным образом. То есть Чэлмерс подчеркивает различие между когнитивной ролью репрезентаций и чего-то, являющегося в них. Исследование первой роли является относительно легким — это просто проблема о природе и функциях репрезентаций. Труднее понять, как в этих состояниях что-то может казаться мне как что-то. Каким образом симуляция ретины глаза, процессы в визуальных областях коры мозга и механизмы различения в итоге дают состояние, когда что-то *явствует* мне определенным образом. Собственно трудность состоит в том, что эта проблема выглядит как *suī generis* и является принципиально иной, нежели проблема о когнитивной стороне сознания, хотя теснейшим образом связана с последней. Если бы удалось найти ответ на эту проблему, это и было бы объяснением сознания. «Сознание является мистическим; когнитивность не является таковой. Онтологически сознание является новизной; когнитивность онтологически является «доступной закуской». Когнитивность может быть объяснена функционально; сознание сопротивляется такому объяснению. Когнитивность целиком управляется законами физики; сознание управляется частично независимыми психофизическими законами»<sup>1</sup>.

Сходную позицию занимает Дж. Сёрль. Характерная черта современных материалистов, говорит он, «боязнь сознания». Они «таким образом переопределяют «сознание», что отрицают главную черту сознания, а именно, его субъективность. Глубочайшей причиной боязни сознания служит то, что сознание обладает этой, в сущности, ужасающей чертой — субъективностью»<sup>2</sup>. Сёрль не призывает повернуться спиной к объек-

<sup>1</sup> Chalmers D.J. The Conscious Mind. In Search of Fundamental Theory. N.Y., Oxford et al., 1996. P. 172.

<sup>2</sup> Сёрль Дж. Открывая сознание заново. М., 2002. С. 69. См. рецензию на эту книгу (англ. изд. 1992 г.): Dennett D. Review on Searle J.R. «The Rediscovery of Consciousness» // Journal of Philosophy. V. XC. № 4. April 1993. P. 193—205.

64

тивному миру науки, однако усматривает ограниченность в нынешнем понимании объективности и даже признаки кризиса науки в том, что та не в состоянии преодолеть асимметрию между научным подходом к объяснению интеллектуальной деятельности и здравым смыслом, опирающимся на субъективный мир ощущений и чувств. Недавно идея, что все физические науки следовало бы революционизировать, с тем чтобы в них нашлось место для чувств и эмоций, была поддержана философом Майклом Локвудом<sup>1</sup>. Более радикально рассуждает Томас Нагель. Он



утверждает, что по мере приближения к субъективному миру сознания, к вопросу, *что это такое* — *быть осознающим бытием*, мы вынуждены оставить за спиной объективный мир науки и обратиться к «точке зрения от первого лица». Даже будущая революционизированная наука не совладевает с такими свойствами. «Субъективные свойства сознательных ментальных процессов — в противоположность их причинам и следствиям — не могут быть схвачены очищенной формой мысли, подходящей для того, чтобы иметь дело с физическим миром, который лежит в основе их появления»<sup>2</sup>.

Убежденный элиминативист в отношении квалиа, Деннет категорически не согласен с Сёрлём, Нагелем, Чэлмерсом. Ни они, ни другие сторонники феноменологии квалиа, говорит он, не в состоянии привести объективные критерии для их идентификации и существования. Все, к чему они апеллируют, — позиция от «первого лица» с ее интуицией и верой в непосредственно очевидное. Их протекционизм в отношении феноменальных качеств достигается «только ценой объявления их вообще находящимися вне какой-либо досягаемости для науки. Если квалиа столь застенчивы, что к ним нет доступа даже с позиции первого лица, тогда вообще невозможна никакая перволичностная наука о квалиа»<sup>3</sup>. Деннет не разделяет

<sup>1</sup> Lockwood M. *Mind, Brain and the Quantum*. Oxford, 1989.

<sup>2</sup> Nagel Th. *The View from Nowhere*. Oxford, 1986. P. 15.

<sup>3</sup> Dennett D. *The Fantasy of First-Person Science* // Internet: <http://ase.tufts.edu/cogstad/imbpage.htm>, P. 7.

65

скептицизма и пессимизма этих философов в отношении объяснения субъективности: дело не в ограниченности науки, а в ограниченности философского воображения, опасющегося неординарных движений мысли. Несмотря на неясности в понимании опыта, несмотря на яростное сопротивление сторонников субъективности (которое, признает он, раньше недооценивал), феноменологии следует бросить «радикальный философский вызов». «Вызов» включает решение двух взаимосвязанных задач: во-первых, показать несостоятельность интроспекции и, во-вторых, показать иллюзорность представления об опыте как состоящем из «квалиа».

Со времен Локка, Юма и Беркли об опыте наговорено предостаточно. В последние десятилетия литература стала необъятной. Учеными детально исследованы свойства нейрофизиологии мозга и внешнего мира вызывать те или иные чувства и ощущения<sup>1</sup>, в их распоряжении появились инструменты, позволяющие наблюдать работу мозга. Много интересных мыслей и гипотез высказано философами. Однако ситуация с опытом не прояснилась, скорее, стала еще более запутанной. И ученым, и психологам, и философам по-разному видится «флора и фауна феноменологического сада», и их картины виденного разительно отличаются.

Видя сумятицу, царящую в изображениях «флоры и фауны феноменологического сада», психологи попытались навести порядок с помощью разного рода классификаций. Обитателей субъективного мира они обычно подразделяют на три группы: 1) опыт «внешнего» мира (зрительные впечатления, звуки, запахи, исчезающие или внезапные чувства, ощущения тепла и холода, чувство положения частей тела); 2) опыт «внутреннего» мира (образы фантазии, внутреннее зрение, сны наяву, разговоры с самим собой, воспоминания, яркие идеи, предчувствия); 3) опыт эмоций или «аффектов» (взрывы гнева, радости, ненависти, замешательства, вожде-

<sup>1</sup> Деннет имеет в виду работы S. Shoemaker (1975, 1984, 1988), S.L. White (1986), P. Kitcher (1979), G. Harman (1990), I. Fox (1989).

66

ления, удивления, порывы гордости, беспокойства, сожаления, иронии, сострадания, благоговения). Деннет считает все эти классификационные упражнения бесполезными: подразделение потока сознания на «внешнее», «внутреннее» и «аффекты» несостоятельно и философски и научно; оно сродни когда-то предпринимавшимся попыткам объединить летучих мышей с птицами или дельфинов с рыбами; такого рода классификация апеллирует к внешней

похожести и сомнительной традиции, нежели к какому-либо глубокому родству феноменов. В конечном счете, считает Деннет, все они запутались в клубке проблем, связанных с *качественными характеристиками опыта*, выпутаться из которого им трудно. Главная причина путаницы и у философов, и у психологов состоит в том, что все они продолжают опираться на интуицию о самом себе и на тезис о достоверности этой интуиции или, говоря словами Р. Рорти, приписывают ей «привилегированный статус». На базе этой интуиции возникло огромное количество взимоподдерживающих доктрин, горизонт которых замкнут стенами Картезианского Театра с его Оком, лицемерным потоком сознания. Несмотря на то, что в его стенах возникает масса парадоксов (парадокс регресса, парадокс никем не означенного Означивателя, парадокс субъективных качеств и др.), и несмотря на несовместимость идеи Театра с новыми научными данными об опыте, они пытаются разрешать парадоксы, не покидая этих стен.

На протяжении многих лет Деннет изучал полученные в когнитивных науках, нейронауках и психологии результаты того, как нам представляются «данные опыта». (Многие приводимые им примеры взяты из наук о мозге и когнитивной психологии, являются очень сложными, понятными только для специалистов, и мы не будем их приводить.) Один из его выводов состоит в том, что психологи и философы, продолжающие держаться интроспективизма и настаивающие на привилегированном доступе человека к его собственному сознанию, слишком доверяют вербальным отчетам испытуемых.

67

В последнее время накоплено большое количество экспериментально подтвержденных фактов, свидетельствующих об ошибочности отчетов испытуемых субъектов о своем ментальном мире. В «Сознании объясненном» он приводит описание многочисленных когнитивных иллюзий. Согласно этим описаниям, испытуемые субъекты часто не в состоянии дать определенные и непротиворечивые ответы, что же они испытывают в опыте, какими именно «данными» они оперируют и каково место их Я в субъективном опыте. Многие из того, что интуитивно им кажется непосредственно воспринимаемым в опыте, на самом деле они вообще не воспринимают, а многое, что они в действительности воспринимают, их Я не видит. Отчеты субъектов о своих внутренних переживаниях включают самокорректировку, зависят от имеющегося у них понятийного багажа и от массы других факторов. Исследователи сталкиваются с фактом, что конкретный опыт, как он представляется субъекту, очень трудно точно фиксировать; его ткань изменяется по мере описания, зависит от времени описания, трансформируется в ходе обучения, взросления и приобретения знания.

Такого рода наблюдения ученых, считает Деннет, делают сомнительной идею о приватном мире, к которому испытуемый имеет безошибочный и привилегированный доступ, а вместе с этим и идею о сознании как Театре или Сцене, наблюдаемой и контролируемой единым режиссером — самостью. Скорее, они свидетельства того, что человеческие существа являются закоренелыми мистификаторами в отношении их собственного чувственного опыта и мышления. Накопленные экспериментаторами результаты говорят о тщетности традиционных попыток эмпириков строить объяснения сознания и знания на основе интуиции о собственных переживаниях. Этим объясняется, почему исследователи в области когнитивных наук, в особенности специалисты по искусственному интеллекту, занимающиеся созданием понимающих язык компьютерных систем, игнорируют «магию феноменологии». Они игнорируют ее в значительной мере

68

«в силу убеждения, что феноменология, какой бы она ни казалась реальной и привлекательной, является нефункциональной — колесом, которое крутится, но не двигает ни одну важную машину восприятия»<sup>1</sup>.

#### 4. «Дисквалификация квалиа»

Философы давно ломали голову над вопросом, что представляют собой онтологически качества нашего опыта. С каким характером бытия мы имеем дело, когда видим, например, красный закат, слышим звук грома, вдыхаем аромат цветка или вкушаем пищу? После того как под напором

физического знания цвета, запахи, звуки, вкусы были изгнаны из «внешнего» мира, современные авторы стали искать для них безопасное место в черепной коробке, приспособив для этого различные имена: «сырые чувства», «sensa», «феноменологические качества», «внутренние свойства сознательных опытов сознания», «качественное содержание ментальных состояний». В последнее время наиболее употребительным стал термин «квалиа». В общей форме можно сказать, что квалиа есть то, каким образом что-то выглядит для нас, кажется нам, в каком качестве оно предстает перед нами. Квалиа не тождественны причинам, их вызвавшим: у боли могут быть объективные физические причины, но ее ощущение — это совсем другое дело. Возникают вопросы: можно ли, не попадая в ловушки субъективного идеализма, полагать квалиа существующими? как мы можем доказать это? каков их статус в экономике когнитивной машины? Одна из глав книги «Сознание объясненное» — «Квалиа дисквалифицированные» — начинается с такого образа. Когда у воздушного шара перепутались стропы, в принципе их можно распутать, имея должное терпение и аналитический

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 56.

69

склад ума. На практике иногда лучше пожертвовать какими-то из них и, отрезав, заменить новыми. Таким Деннету видится подход к философской проблеме квалиа — «мучительному спутанному клубку все более растущих и закручивающихся вихрями мысленных экспериментов, жаргонов, in-jokes, аллюзий ... и попусту потраченных усилий»<sup>1</sup>.

Для пояснения того, почему нам кажется, что нечто ментальное и феноменальное существует, хотя на самом деле ничего такого нет в природе и при строгом подходе все это следует элиминировать из теоретического языка, Деннет предложил такой мысленный пример.

Предположим, что где-то мы обнаружили общество, в котором отсутствует знание человеческой физиологии, но которое говорит почти на таком же языке, как английский, за исключением одного курьезного семейства идиом. Когда члены этого общества чувствуют себя усталыми, они говорят, что их осаждают *усталости* (fatigues), что у них имеются ментальные усталости, мышечные усталости, усталости в глазах и усталости в духе. В их спортивном языке фигурируют такие максимы: «слишком много усталостей помешали достижению нашей цели» и «пять усталостей в ногах стоят десяти усталостей в руках». Когда мы входим в контакт с ними и рассказываем им о нашей науке, они спрашивают, *что же представляют собой усталости*. Их интересует, являются ли усталости, которые приходят, уходят и возвращаются, нумерически теми же самыми; есть ли у них своя определенная локация в материи, пространстве и времени, идентичны ли они некоторым специфическим физическим состояниям и процессам или событиям в их телах, или же они состоят из особого рода вещества. Понимая, что у членов этого сообщества изначально неверный подход к этим вопросам, мы могли бы разъяснить им, что таких вещей, как усталости, просто не существует, что у них — путаная онтология. На что они могут возразить: «Вы полагаете, что таких вещей, как усталос-

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 369.

70

ти, не существует? Пробежите несколько раз вокруг дома, и вы лучше убедитесь, так ли это! Есть много вещей, которым ваша наука могла бы научить нас, но несуществование усталостей не относится к ним».

Конечно, это общество вправе говорить об усталостях. В конце концов, это их язык. Можно даже согласиться считать по меньшей мере некоторые их заявления относительно усталостей истинными или ложными в зависимости от того, находятся ли индивиды в дремоте, в изнеможении или только притворяются, можно еще привести физиологические условия истинности или ложности этих заявлений. Однако никак нельзя принимать всерьез онтологию этих заявлений и любую попытку *идентификации* усталостей. Здесь мы должны быть элиминативными материалистами, ибо признаем вне закона существование смутных элементов. Усталости не являются хорошими теоретическими сущностями, как бы ни был укоренен термин «усталость» в «фолк-психологии» воображаемого общества.

Сказанное об «усталостях» воображаемого народа, говорит Деннет, истинно по отношению к нашим собственным верованиям, желаниям, болям, ментальным образам, опытам, квалиа — в том смысле, в котором они понимаются в *обычной жизни*. Хотя, замечает он, «убедить кого-либо, что болей и верований не существует, столь же трудно, как и убедить наш воображаемый народ, что не существует таких вещей, как усталость»<sup>1</sup>.

Одним из пионеров, попытавшихся разобраться в том, что собой представляет чувственный опыт, был Джон Локк. Он отнес цвета, запахи, звуки, вкусы к вторичным качествам, а размеры, формы, движение и твердость — к первичным. Локк не считал вторичные качества «вещами в сознании», скорее видел в них силы или диспозиции объектов внешнего мира провоцировать в голове человека определенные свойства. Однако еще при его жизни возник коварный вопрос: где на-

<sup>1</sup> Dennett D. *Brainstorms*. P. XX.

71

ходятся вторичные качества при отсутствии воспринимающего субъекта? Для своего времени, говорит Деннет, локковская теория имела свои достоинства, ибо вторичные качества были «отшпилены» от внешних объектов. Однако получалось так, что они оказались «пришпиленными» к сознанию в виде *идей* цвета, вкуса и т.д. В этом случае возникает следующий коварный вопрос: какого рода бытием, например, является идея красного цвета? Является ли она окрашенной или бесцветной? В конце концов, что такое «красный цвет»? Где он находится — в мозгу или во внешнем мире?

Несмотря на то, что после Локка была основательно исследована физика цвета и была доказана зависимость ощущения цвета от излучения объектами световых волн определенной длины, ясного ответа на эти вопросы, считает Деннет, до сих пор не получено. В современном стандартном учебнике о мозге дается, например, такое определение цвета: «В мире не существует «цвета» как такового; он существует только в глазах и мозгу зрителя. Объект отражает множество световых волн разной длины. Однако сами по себе эти волны не имеют цвета»<sup>1</sup>. В нем содержится явное противоречие. С одной стороны, авторы утверждают, что цвет существует не «в мире», а в глазах и мозгу зрителя. Однако и мозг и глаза находятся в том же самом физическом мире, являются его частью. Как и все объекты этого мира, они имеют цвета: глаза могут быть голубыми, коричневыми, а мозг имеет разные цветовые оттенки. Заменяя цвет волнами различной длины, ученые вроде бы оторвали его от внешнего мира, но, с другой стороны, они загнали его в «глаза» и «мозг». Где и в каком виде он там прячется? Путаница существует не только относительно субъективного, но и объективного местоположения цвета. Очень трудно сказать, к каким именно свойствам вещей в мире можно отнести цвета. Оказалось, что простая и привлекательная идея, ассоциирующая цвет с уникальными световыми волнами определенной длины и рассматривающая «красноту» как свойство

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 370.

72

отражения всех красных световых волн определенной длины и поглощения всех иных световых волн, не является верной. Согласно экспериментам, поверхности с существенно разными отражательными способностями могут восприниматься, как имеющие один и тот же цвет, а одна и та же поверхность при разном освещении может восприниматься в разных цветах. Фиксированного свойства поверхности, всегда дающего красный цвет, не существует. Иначе говоря, попадающие в глаз световые волны только косвенным образом связаны с цветом, в каком нам видятся объекты<sup>1</sup>.

Все эти научные изыскания, говорит Деннет, можно истолковать как сплошные загадки и даже как вводящий в депрессию факт; напрашивается пессимистическая мысль, что наше представление о перцептуальном схватывании мира выглядит значительно хуже, чем мы думали; получается так, что мы живем в мире сновидений или являемся жертвами массовой иллюзии. Почему так?

Разумеется, если принять тезис сторонников субъективной реальности о существовании чего-то внутреннего, ментального, эпифеноменального, обладающего таинственной способностью феноменологически окрашивать вещи, вроде бы создается видимость объяснения цвета. Однако

при этом невозможно освободиться от вопроса об источнике этой способности. Откуда она? С другой стороны, если свести цвет только к волнам определенной длины, излучаемым объектами, как это нередко делается, цвет вообще исчезает! «Это давняя философская головоломка, которой сегодня мы должны взглянуть в лицо»<sup>2</sup>. Посмотрим, как с ней справляется Деннет и какую конструктивную альтернативу субъективизму и феноменализму он предлагает.

Резон, побуждающий философов принимать квалиа за нечто существующее, говорит он, прост: если наука показала, что цве-

<sup>1</sup> Деннет ссылается на работы по зрительному восприятию следующих авторов: Goras, 1984; Hilbert, 1987; Hardin, 1988.

<sup>2</sup> Dennett D. *Consciousness Explained*. P. 371.

### 73

та не могут находиться вовне, то есть в физическом мире, значит, они должны находиться внутри, а то, что находится внутри, не может быть просто бесцветной идеей или суждением об окрашенных объектах, оно само должно быть каким-то качеством — квалиа. Несмотря на кажущуюся очевидность, данный резон ошибочен. Поскольку «вторичные качества» чрезвычайно резистентны для прямолинейного и четкого (философского) определения, Деннет предлагает для начала обратиться к тому, что говорит наука. Все, о чем она свидетельствует, состоит в следующем: свойства объектов, отражающих свет, являются причинами, которые обуславливают включение в мозг различных дискриминационных состояний, которые, в свою очередь, задействуют лежащие в их основе внутренне эволюционно встроенные диспозиции, а вместе с ними приобретенные диспозиции и привычки разного уровня сложности. А как насчет самих этих свойств? У дискриминационных состояний мозга имеются различные «первичные» свойства (механические, соединительные, состояния возбуждения элементов и др.), благодаря которым включаются «вторичные» или диспозиционные свойства.

У владеющих языком человеческих существ эти дискриминационные состояния и диспозиционные свойства могут задействовать диспозиции к вербальным суждениям в отношении «цвета» различных вещей. Когда кто-либо произносит фразу «Я знаю, что круг на самом деле не является розовым, хотя мне он действительно видится розовым», первая часть предложения выражает суждение о чем-то в мире, а вторая часть предложения выражает суждение второго порядка относительно дискриминационных состояний о чем-то в мире. Может возникнуть предположение: нет ли у наших дискриминационных состояний еще и особых, добавочных, «внутренних» свойств — субъективных, частных, неопределенных, — конституирующих способы, какими вещи выглядят для нас (звуки для нас, запах для нас и т.д.). Если да, то их можно было бы посчитать кандидатами для реальности квалиа. Таких внутренних свойств, утверждает Деннет, нет. Их невозможно обнаружить никакими объективными методами.

### 74

Все, что связано с цветом, является, в конце концов, свойствами «out there», которые поддаются объективному исследованию без необходимости заглядывания «вовнутрь».

Таким образом, на место локковской «идеи красного» Деннет ставит понятие «дискриминирующее состояние», имеющее содержание «красное» на основе врожденных и приобретенных диспозиций. Смысл понятия «дискриминирующее состояние» поясняется им на примере сравнения оттенков цветов человеком и машиной. Сравнить цвета вещей можно по-разному: расположив их рядом и вынося суждения об оттенках, а можно сравнивать цвета в воображении или по памяти — в некоем хранилище ментальных образов. Когда перед нами ставится задача определить, является ли стандартный красный цвет на американском флаге таким же, как на одежде Санта Клауса, в нашем сознании что-то происходит, в результате чего информация извлекается из памяти. Происходящий в мозгу процесс в чем-то аналогичен тому, который может происходить в машине — некоем роботе, приспособленном делать такие сравнения. Например, в робот Carblinind Mark 1 была заложена информация по распознаванию цветов. Когда роботу дается задание распознать степень различия красного цвета на американском флаге и на одежде Санта Клауса, он выдает ответ в цифровом измерении, например, первый



красный цвет #163, а второй #172, то есть разница составляет # 9 единиц. «Это значит, что робот не обращается к фантазии, чтобы привести в действие память (или его настоящее восприятие), но таким же образом поступаем и мы. Робот, вероятно, не знает, как он сравнивает цвета, то есть чего-то видимого ему, с чем-то, что он помнит, но точно так же поступаем и мы»<sup>1</sup>.

Конечно, существует огромное различие в сложностях диспозиций, на основе которых эту задачу выполняет Carblinind Mark 1, и как это делаем мы, но оно не является качественным. Поэтому вполне оправдан вывод, что подобно тому, как Carblinind Mark 1 не нуждается ни в каких добавочных, субъек-

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 374.

75

тивных, ментальных квалиа для выполнения своей задачи, мы тоже не нуждаемся в этих добавках. «Я утверждаю, что у нас вообще нет никаких квалиа. Тот вид различия, который, как полагают люди, существует между любой машиной и любым имеющим опыт человеческим бытием, .. я решительно отвергаю: такого вида различия не существует. Во всяком случае, так это мне представляется»<sup>1</sup>.

Ситуация с цветом немного проясняется, если мы зададимся вопросом: «Почему существуют цвета и для чего они нужны?» Ответ на него следует ждать от науки, однако философские спекуляции здесь весьма уместны. Интересную мысль высказала философ нейронауки Кэтлин Экинс (1989, 1990). Суть ее в том, что цвета и цветовое восприятие были созданы эволюцией друг для друга. Схожую мысль развивает психолог Н. Хэмфри: цветовое видение существ с самого начала коэволюционировало с цветами предметов, чье *raison d'être* было кодирование цвета. Цветовое видение насекомых развивалось вместе с цветовыми различиями растений, которые они населяли. Без цветовых кодов растений и кодов цветовых оттенков у них не было бы развито цветовое видение, и наоборот: без наличия излучения объектами волн определенной длины не возникло бы цветовое кодирование. Поэтому принцип цветового кодирования является основой цветового видения у насекомых. Одни цвета (запахи, звуки) были созданы для того, чтобы ими наслаждаться, другие — чтобы их избегать, и то и другое чрезвычайно важно для их выживания. Деннет считает философски очень важным вывод, что «с самого начала цвета были созданы для того, чтобы быть видимыми теми, кто был создан для того, чтобы видеть их»<sup>2</sup>. Диспозиция к распознаванию цветов, говорит он, как и диспозиция к любви, развивалась при наличии двух сторон. То же самое можно сказать об эволюции цветового видения у других видов живых существ. Хотя здесь остается еще много загадок. Не у

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 375.

<sup>2</sup> Ibid. P. 378.

76

всех живых существ, имеющих глаза, есть любой вид цветового видения. Птицы, рыбы, насекомые и рептилии обладают цветовым видением, чем-то вроде нашей «трехцветной» (красное—зеленое—синее) системы; у собак и кошек она отсутствует. Среди млекопитающих только приматы имеют цветовое видение, однако и между ними наблюдается поразительное различие. Почему у одних видов возникло цветовое видение, а у других нет, — это еще предстоит узнать науке. Пока знание о вторичных качествах остается со многими белыми пятнами. «Все это оказывается вдохновляющей и сложной историей, все еще в значительной степени спекулятивной»<sup>1</sup>. Можно предположить, что возникшее в коэволюции цветовое зрение живых существ затем было использовано ими по отношению к неорганической материи. Восприятие зеленого яблока сделало возможным восприятие зеленых изумрудов.

Так или иначе, ошибочно думать, говорит Деннет, что сначала были цвета — цветные горы, цветное небо, красно-оранжевая ржавчина и ярко-синий кобальт, — и только потом Мать-Природа двинулась дальше и стала использовать эти свойства для цветового кодирования вещей и восприятия кодов. «Скорее, изначально существовали различные отражательные свойства поверхностей, реактивные свойства фотопигментов и т.д., а Мать-Природа развила из этих сырых материалов эффективные взаимно приспособленные системы «цветового» кодирования /

«цветового» видения («color»-coding / «color»-vision systems), и среди свойств, возникших из этого процесса дизайна, оказались свойства, которые мы, нормальные человеческие существа, называем цветом»<sup>2</sup>. Восприятие синего кобальта и синего крыла бабочки как синего цвета есть побочный продукт исходного процесса, породившего цветное зрение в неразрывном единстве с цветовым кодированием.

С гипотезой Деннета о диалектике цветового кодирования мало кто спорит. Оспаривается «дисквалификация квалиа». Кри-

<sup>1</sup> Ibid. P. 377.

<sup>2</sup> Ibid. P. 378.

77

тики говорят, что никакое эволюционное объяснение механизмов возникновения цвета не снимает онтологического вопроса о «бытии цвета» — что же он такое? То есть физикалисту следует справиться с *фактом* эпифеноменальности квалиа. Даже если физикализм будет стремительно прогрессировать и идеальная наука будет обладать полнотой физикалистского знания, говорят они, квалиа, расцвечивающие наш опыт, выпадут из него.

В 1982 году Фрэнк Джексон в статье «Эпифеноменальные квалиа» предложил мысленный эксперимент, нацеленный показать уязвимость физикалистского «аргумента полноты знания» в споре с защитниками эпифеноменальности квалиа<sup>1</sup>. (На протяжении последующих двадцати лет этот эксперимент был предметом толкования с самых различных позиций, и из него делались порою противоположные выводы.)

Им придумана следующая история про Мэри. Мэри является блестящим ученым, но по каким-то непонятным причинам вынуждена заниматься исследованием мира, сидя в черно-белой комнате и наблюдая его через черно-белый телевизионный монитор. Будучи специалистом по нейрофизиологии зрения, она обладает всей полнотой физической информации, которой только можно обладать относительно того, что происходит, когда мы видим красные помидоры, голубое небо и т.д. Возникает вопрос, что произойдет, когда Мэри высвободится из ее черно-белого заточения или ей дадут цветной телевизор? Увидев другие цвета, научится ли она чему-нибудь дополнительно или нет? Для Джексона очевидно, что она узнает что-то новое о мире и человеческом визуальном опыте, и поэтому ее предшествующий опыт следует признать неполным, даже при условии наличия у нее *всей полноты* физической информации о мире. Его вывод таков: «Ergo, существует большее, что может находиться в наличии, нежели физическая информация, следовательно, физикализм ложен...»<sup>2</sup>.

<sup>1</sup> Jackson F. Epiphenomenal Qualia // *Philosophical Quarterly*. V. 32. 1982. P. 127—136.

<sup>2</sup> См: Dennett D. *Consciousness Explained*. P. 398.

78

Очень многие люди, в том числе закаленные в боях дискуссий философы, говорит Деннет, сочли этот мысленный эксперимент убедительным аргументом, то есть согласились, что помимо знания о фактах существует еще нечто дополнительное, и приняли вывод, что «зрительный опыт имеет квалиа, которые "эпифеноменальны"»<sup>1</sup>. Они не заметили в нем существенных изъянов. Не обратили внимание на то, что критическая посылка эксперимента состоит в том, что Мэри располагает *всей полнотой* физической информации, а не просто имеющейся в наличии информацией. Если исходить из такой посылки, тогда выводы эксперимента могут быть другими, то есть в пользу физикализма.

Чтобы это увидеть, Деннет предложил продолжить мысленный эксперимент с Мэри-ученой. «И вот однажды люди, заточившие Мэри в черно-белую комнату, решили, что для нее настало время увидеть цвета. Но они прибегли к трюку и в качестве первого цветового опыта дали ей синий банан. Взглянув на него, она сказала: «Вы что, пытаетесь сыграть со мной шутку? Бананы желтые, а этот синий». Заточившие ее люди были обескуражены. Как она догадалась? Очень просто, ответила она. «Вы должны помнить, что я знаю все на свете, абсолютно все, что когда-либо было известно о физических причинах и следствиях цветового зрения. И, конечно, перед тем, как вы принесли банан, я уже описала во всех мельчайших деталях, какое именно физическое

воздействие оказывает желтый (синий, зеленый) объект на мою нервную систему. Поэтому мне досконально известно, какие именно мысли у меня возникнут (потому что, в конце концов, «простая диспозиция» думать относительно того или иного не является одной из ваших знаменитых квалиа, не так ли)... Я понимаю, что вам трудно вообразить огромность моего знания о собственных реактивных диспозициях и пред-

<sup>1</sup> Например, М. Тье, 1986; D. Lewis, 1988; B. Loar, 1990; W.G. Lycan, 1990; L. Nemirov, 1990; G. Harman, 1990; N. Block, 1990; van Gulick, 1990. Единственный, кто попытался сопротивляться такому выводу, был Пол Черчленд.

79

ставить, что способ, каким синий цвет воздействовал на меня, не явился для меня сюрпризом. Конечно, вам трудно вообразить последствия того, что кто-то обладает абсолютно всем физическим знанием о любой вещи»<sup>1</sup>.

Представить человека, который знает все на свете о физических каузальных факторах зрения, говорит Деннет, конечно, трудно, но логически помыслить возможно, и отметить эту возможность для чистоты эксперимента нельзя. Авторы, использующие мысленный эксперимент Джексона в качестве аргумента в пользу эпифеноменальности цвета, упускают из виду, помимо «аргумента от полноты знания», еще один аргумент: заточенная в черно-белую комнату Мэри имела опыт распознавания черного и белого цветов и, конечно, оттенков серого, следовательно, у нее была сформирована *диспозиция к распознаванию цветов*, и на ее основе ей не составило труда отличить синее от желтого.

Общий вывод Деннета таков: эпифеноменализм — прямой путь к радикальному солипсизму, отрезающему сознание и опыт от коммуникации с материальным миром: получается, что квалиа существуют, активно окрашивают наш опыт и в то же время не оказывают никакого эффекта на физический мир. Сторонники существования квалиа принимают их за «факты» при посылке, что их нельзя исследовать никакими годящимися для физического мира способами. Объективность научного знания ставится под угрозу.

## 5. Иметь опыт — значит иметь суждение об опыте

Итак, для того чтобы сохранить нашу веру в объективность мира и в способность науки давать объективное знание, мы должны отвергнуть идею о существовании квалиа как некоей добавки к миру. Что же тогда присутствует в

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 399—400.

80

опыте, когда мы видим красную розу или слышим шум моря? Все многообразие нашего опыта, отвечает Деннет, конституируют наши суждения — лингвистические параметры опыта. Притягательность Картезианского Театра объясняется тем, что субъективно он кажется нам естественным местом, где к тому, что нам кажется на основе интуитивного самоощущения феноменов опыта, добавляются суждения о наличии этих феноменов. Однако это самообман. Философы, которые помимо суждений постулируют еще и «реальные кажимости», в действительности занимаются умножением сущностей без необходимости.

В ключевом пассаже «Сознания объясненного», обсуждая неоновый цветораспределительный феномен (в котором цветовой круг видится существующим в белом пространстве между окрашенными неоновыми линиями), Деннет придумал следующий диалог с «Отто» — воображаемым критиком, выступающим от имени здравого смысла.

«Д е н н е т: Такой вещи, как розовый круг, не существует, он просто кажется нам существующим. Отто: Послушай, я говорю не только, что мне кажется, будто здесь существует яркий розовый круг; здесь реально видится наличие яркого розового круга!

Деннет: Спешу согласиться с тобой. Я никогда не обвинял тебя в неправдивости. Ты действительно означаешь его, когда говоришь, что розовый круг тебе кажется существующим.

Отто: Вникни в то, что я говорю. Я не просто означаю его. Я не только думаю, что мне кажется, что здесь имеется яркий розовый круг; мне реально видится наличие яркого розового круга!

Деннет: Сейчас ты сделал его. Как и многие другие, ты попал в ловушку. Ты, наверное, думаешь, что существует различие между думанием (суждением, решением, искренним мнением о чем-то), что нечто выглядит розовым для тебя, и чем-то реально выглядящим розовым для тебя. Однако здесь нет различия. Не существует такого феномена, как реально

81

выглядящего, помимо и над феноменом того или иного суждения, что имеет место то-то и то-то»<sup>1</sup>. Разумеется, Деннет не внушает Отто, что у того нет ощущения розового круга, он убеждает его в другом — в его мозгу не существует ментального качества «розоватости». Его принципиальный тезис состоит в том, что между субъективным ощущением, что нечто выглядит розовым, и суждением, что это «нечто» реально выглядит розовым, нет различия. То есть отрицается различие между воспринимаемым в опыте феноменом и мыслью об этом феномене. А мысль, естественно, оформлена лингвистически. «Нормально достаточным условием для того, чтобы иметь опыт чего-то, является последующий вербальный отчет»<sup>2</sup>. Все, что мы принимаем в опыте за реальности, в действительности только кажимости, реально существуют только отчеты, верования, мнения, суждения.

Поэтому предметом теории сознания должны *стать пропозициональные эпизоды* нашего опыта и суждения о сознании. В уже упоминавшейся работе «Об отсутствии феноменологии» (1979) Деннет пишет: «Я хочу защитить точку зрения, что такие суждения исчерпывают наше непосредственное сознание, что в нашем индивидуальном потоке сознания нет ничего, кроме таких пропозициональных эпизодов. ...Грубо говоря, она сводится к тому, что не существует феноменального двойника любому из таких отношений к нашим отчетам. В наличии имеются производимые нами публичные отчеты, затем есть эпизоды нашего пропозиционального осознания, наши суждения, а за всем этим простирается — в той мере, в какой имеется в виду интроспекция, — темнота»<sup>3</sup>.

Здесь важно иметь в виду следующее. Когда Деннет говорит, что кажимости нашего опыта — это только пропозициональные отношения, он не повторяет Витгенштейна. Его когнитивизм не является чисто лингвистическим, его правильнее назвать эволю-

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 363—364.

<sup>2</sup> Ibid. P. 140.

<sup>3</sup> Dennett D. *On the Absence of Phenomenology*. P. 95.

82

ционным когнитивизмом. Везде и всюду он стремится подвести под когнитивизм натуралистические (эволюционные) подпорки. Его важнейшая посылка состоит в том, что, взятое в изоляции, ни одно свойство сознания нельзя считать проявлением какого-либо качества, все свойства являются результатом сложнейших процессов природной и культурной эволюции. Мы ничего не проясним, если будем принимать<sup>1</sup> за критерии отдельные свойства нашей психологии и обыденных представлений. За всеми ими скрываются природные диспозиции.

Из сказанного видно, что деннетовская онтология совершенно иная, нежели картезианская. Его модель является не альтернативным объяснением тех же самых феноменов, которые присутствуют в модели Картезианского Театра, а скорее их элиминацией. Если в картезианской модели «Око» наблюдает поток феноменов и на их основе выносит суждение, в деннетовской модели опыт состоит из природных диспозиций и суждений. Главный акцент здесь сделан на лингвистическом, когнитивном аспекте сознательных процессов. В картезианской модели, если мне кажется, что я слышу гром или воображаю зеленого тигра, тогда я действительно слышу его и вижу тигра, пусть не во внешнем мире, а в моем приватном театре воображения; если мне кажется, что Я являюсь единой самостью, следовательно, мое единое Я существует. В деннетовской онтологии нет места ни для существования субъективных феноменов и образов, ни самости, какой она представляется мне. Опыт сводится к суждениям, верованиям в то, что имеет место некоторое состояние сознания и эти суждения и верования могут быть объективно (интерсубъективно) проверены. Опыт боли становится болью при условии отчета индивида о ней, который эмпирически проверяем на основе исследования неполадок организма. Ведь субъективная боль может быть и фантомом. То, что

считается болью, слуховым ощущением, моим Я и т.д., получает смысл только при оформлении в языковые выражения «мне больно», «я слышу звук», «я являюсь человеком».

Одним словом, по Деннету, наши верования в существование «квалиа» получают смысл на когнитивном, информацион-

83

ном уровне. Люди являются социальными самостями, и многое в их ментальных способностях, включая ощущаемые качества опыта и эмоции, встроено через коммуникации с другими людьми и через обучение языку. Конечно, если то, из чего состоит человеческое сознание, является преимущественно производством лингвистически оформленных верований и суждений о них самих и их окружении, тогда восприятие человеком его внутренней субъективной жизни будет иллюзорным.

В суммарном виде негативный тезис Деннета в отношении феноменологической стратегии выражен в «Сознании объясненном» следующим образом. «Самой естественной вещью в мире является представление о сознании, действующем в чем-то вроде Картезианского Театра... [Но] стоит нам со всей серьезностью посмотреть на происходящее за его кулисами (в науках о мозге, в эмпирической психологии), как обнаруживается, что в действительности мы не видим того, что, как нам казалось, мы видим на его сцене. Огромная пропасть между феноменологией и физиологией немного сужается; мы видим, что некоторые «очевидные» черты феноменологии вообще не являются реальными: не существует заполняемости воображения; не существует внутренних квалиа; не существует центрального источника значения и действия, не существует магического места, где происходит понимание. Картезианского Театра в действительности не существует...»<sup>1</sup>.

## 6. Почему нельзя создать чувствующий боль компьютер?

Итак, по Деннету, квалиа, которые окрашивают наш опыт, — ощущения и представления, образы — это только кажимости (Seemings). В реальности имеются природные диспозиции и лингвистическая интерпретация состояний опыта. Но как быть

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 434.

84

с болью? Когда молоток ударяет по пальцу, боль возникает без всякой рефлексии и интерпретации. Этот факт — твердый орешек как для феноменалистов, так и антифеноменалистов.

В англоязычной литературе за последние полвека опубликовано огромное количество работ по проблеме боли. (Предмет удивления российских философов.) Проблема обсуждается в философии сознания, эпистемологии, психологии, теориях искусственного интеллекта, когнитивных науках. Повышенное внимание объясняется прежде всего тем, что феномен боли оказался очень удобным для анализа опыта с целью прояснения его перцептуальных и рациональных сторон, а вместе с этим и сознания. А также тем, что пример с болью позволяет лучше понять возможности (и границы) компьютеров.

Природа ощущений, в том числе природа боли, как известно, занимала еще эмпириков XVII века. В нынешней аналитической философии этот вопрос приобрел множество новых смыслов и неожиданных оттенков, не говоря уже о решениях. Обсуждаются, например, такие вещи: можно ли говорить о боли как «факте», а если да, что можно считать «фактом боли» — особое ментальное качество сознания, или нейрофизиологическое состояние организма, или лингвистически оформленные реакции, или болевое поведение организма? Бывает ли неощущаемая боль или боль, которая не проходит через сознание? Можно ли заблуждаться относительно того, есть у тебя боль или нет, или ложно определять место боли? Является ли боль только «приватным» ощущением? Возможно ли чувствовать чужую боль? Могут ли два разных человека чувствовать одну и ту же боль? Можно ли говорить о боли применительно к животным? Может ли компьютер чувствовать боль? В конечном счете все эти вопросы сводятся к главному: существует ли *в реальности* особое ментальное качество, которое мы обобщенно именуем



«болью»?

В идущей от английских эмпириков традиции принято считать, что в описании опыта в обычном словоупотреблении (или, как ее сегодня называют, фолк-психологии) действует юмовский принцип *esse est percipi*. Данный принцип предпо-

85

лагает три вещи: 1) каждый из нас *непосредственно* знает, находится ли он в состоянии боли или нет; 2) непосредственное знание является абсолютно *достоверным*; 3) подтверждением достоверности является *интроспекция* внутренних ментальных состояний. В отличие от знания собственной боли, знание о боли других людей является выводимым и менее достоверным. Общий гносеологический вывод из этой традиционной схемы опыта состоит в том, что существует совокупность простых, не поддающихся дальнейшему анализу, относящихся к опыту предикатов, которые могут служить основанием для определения всех сложных понятий, составляющих сознание.

Неудовлетворительность этой схемы была замечена задолго до появления аналитической философии. В XX веке наиболее радикальную попытку ее поломать предприняли бихевиористы, сведя боль к ее внешним причинам и реакциям на них организма. В отличие от классического бихевиоризма, сводящего «боль» к болевому поведению организма, Витгенштейн утверждал: оправданием утверждения, что кто-то находится в состоянии боли, является болевое поведение *в обстоятельствах*, вызывающих боль.

В ходе дискуссий об идеях Витгенштейна многими философами была поставлена под сомнение естественная интуиция человека, убежденного, что субъективное ощущение боли свидетельствует о существовании у него особого ментального состояния — боли. Эту скептическую линию продолжил Деннет. Как уже говорилось, существенную роль во взглядах Деннета на феноменальный аспект сознания играет инструмент, которого еще не было во времена Витгенштейна, а именно — компьютерные аналогии. В последние десятилетия достигнут существенный прогресс в симуляции компьютером разумного логического мышления, однако он не сопровождался таким же прогрессом в симуляции ощущающей деятельности. Созданы компьютеры, способные играть в шахматы, решать логические и математические задачи, предсказывать новые химические соединения, делать переводы, вести разум-

86

ный разговор на естественном языке, писать детективные истории, сочинять музыку и даже радоваться хорошей погоде. Тем не менее создать *чувствующие боль* компьютеры не удалось. Данный их изъян дал философам богатую пищу для размышления о причинах такого несоответствия. В чем причина диспропорции между успешной реализацией в компьютерах высокоинтеллектуальной когнитивной деятельности и неуспехом в реализации в них ощущений и эмоций?

На этот счет были высказаны разные суждения — крайне пессимистические, умеренно пессимистические и оптимистические. Крайне пессимистическую (негативистскую) позицию занимает, например, Томас Нагель. Поскольку никакой компьютер не сможет дублировать всегда специфические внутренние переживания и ощущения человека, компьютерная (физикалистская) стратегия заведомо обречена на неудачу. «Подлинные принципы, лежащие в основе сознания, могут быть раскрыты, если вообще будут раскрыты, прямым доступом к нему»<sup>1</sup>. Он вообще сомневается в нужности и правомерности компьютерных гипотез для понимания сознания и познания. Все попытки философов понять чувствующее и осознающее себя бытие по аналогии с искусственно созданными компьютерами, с его точки зрения, — это «гигантское и впустую потраченное время»<sup>2</sup>. Так думают далеко не все. Значительно большее их число высказываются за «слабый AI». Компьютерные метафоры, считают они, придали новое дыхание философии сознания, расширили горизонты ее воображения и позволили понять многое в функционировании сознания, что без них было бы невозможно. Джонатан Коэн (Оксфорд), например, видит большие достоинства компютационной гипотезы в прояснении формальных структур мышления. Вместе с тем он полагает, что чересчур расширительно понимаемая компютационная гипотеза неспособна

разрешить все проблемы, относящиеся к ментальным состояниям, которыми занимались философы от Декарта до наших

<sup>1</sup> Nagel Th. *The View from Nowhere*. N.Y., Oxford. P. 16.

<sup>2</sup> *Ibidem*.

87

дней. «Такие проблемы принципиальным образом относятся к состояниям сознания, которые характеризуются перцептуальным или воображаемым опытом, или знакомыми чувствами убежденности, радости, печали, страха, наслаждения и другими чувствами»<sup>1</sup>. У оптимистически настроенных философов — сторонников «сильного AI» — другое мнение. Неудачи компьютерной симуляции боли и других аффектов, считают они, носят временный характер и будут преодолены по мере совершенствования и «мягких» и «жестких» программ. Деннет, разумеется, находится в стане оптимистов и сторонников «сильного AI». Тем не менее у него особое мнение относительно неуспехов симулирования компьютером феномена чувствительности, в частности боли. Рассмотрим его.

В 1971 году была опубликована книга Кейта Гундерсона «Ментальность и машины», вызвавшая большие дебаты<sup>2</sup>. В ней он поднял вопрос о причинах «упрямого нежелания» компьютерных машин симулировать чувствительность. Главная из них, с его точки зрения, состоит в недооценке разработчиками различия между программно-рецептивными и программно-резистентными чертами ментальности. Успехи компьютеров были связаны с симулированием целеориентированных или *sapient* черт ментальности, крушения — с *sentient* или чувствительными (животными) чертами ментальности. Мы можем заставить машину *думать*, но не можем заставить ее *чувствовать* по той причине, что вторая задача является *принципиально иной*, нежели первая. Предположение Гундерсона свелось к следующему: решение второй задачи зависит не столько от софистичных инноваций в *программном обеспечении*, то есть создании «мягких программ» нового поколения (software), сколько от разработки принципиально новых видов *жестких программ* компьютеров (hardware), если такое решение вообще возможно.

<sup>1</sup> Cohen L.J. *The Dialogue of Reason. An Analysis of Analytical Philosophy*. Oxford, 1986. P. 194.

<sup>2</sup> Gunderson K. *Mentality and Machines*. 1971.

88

Деннет откликнулся на работу Гундерсона статьей «Почему мы не можем создать чувствующий боль компьютер» (1978)<sup>1</sup>. В ней он в чем-то согласился с Гундерсоном, однако его главную посылку о наличии программно-рецептивных и программно-резистентных черт ментальности счел неубедительной, а предложенную им перспективу — ведущей в ложном направлении. Рассмотрим основные его аргументы.

Большинство как оптимистических, так и пессимистических рассуждений о компьютерах, способных или неспособных симулировать боль, говорит он, характеризуется принятием априорной установки, что боль является особым ментальным качеством живого существа. То есть принимается тезис, который еще не доказан. Лучший способ выявить и продемонстрировать априоризм — обращение к деталям и проведение философских мысленных экспериментов.

В качестве предмета экспериментальной проверки Деннет взял два предположения: а) создана реальная мягкая программа боли для компьютера (software); б) создан принципиально новый дизайн ощущающего боль робота или жесткая программа (hardware). Проанализировав множество логических ходов в защиту этих предположений (мы не имеем возможности их воспроизводить, они заняли бы слишком много места), он сделал вывод, что и в случае (а), и в случае (б) возникают *неразрешимые логические трудности*, что свидетельствует о необходимости принять принципиально иное направление поиска.

Противоречия, возникающие при создании чувствующего боль компьютера, по Деннету, являются следствием принятия априорной посылки, что боль является особым ментальным качеством опыта. Исходя из нее, ставится задача поиска программы, дублирующей это ментальное качество. На самом деле первым шагом поиска должно быть рассмотрение смысла этого понятия с точки зрения правомерности того, что понимается

<sup>1</sup> Она была первоначально опубликована в журнале «Synthese» и включена в книгу «Головоломки»: Dennett D. Why You Can't Make a Computer that Feels Pain // Dennett D. Brainstorms. Ch. 11. P. 190—229.

89

под «болью». Анализ показывает, что в повседневном языке понятие «боль» употребляется разноречиво и противоречиво. Для подтверждения этой констатации он предлагает обратиться к одному из широко дебатированных вопросов «может ли существовать боль, которую не чувствуют?» и посмотреть, насколько однозначно используются там смыслы слова «боль». Им предлагается такой пример. Известно, что морфий делает организм нечувствительным к боли. При принятии же больших доз этого препарата появляется боль, предупреждающая организм об отравлении. Если считать боль своего рода противоядием против отравления морфием, тогда про морфий нельзя сказать, что он помогает подавлять болевые ощущения. По Деннету, справиться с этим традиционным вопросом можно только путем его дискредитации.

В обыденной речи, говорит Деннет, мы часто псевдообъекты принимаем за реальные объекты. Именно так обстоит дело с вопросом о «существовании» боли. Например, я чувствую зубную боль, но во время эмоционально захватившего меня разговора с приятелем она прошла и тотчас же возобновилась, как только закончился разговор. Физические каузальные причины боли действовали все это время. Однако куда делась «боль», если ее считать ментальной сущностью, во время разговора? Где она пряталась все это время? Можно ли вообще говорить о ее онтологическом существовании, если она не ощущалась?

Эти простые примеры, считает Деннет, свидетельствуют о неоднозначном употреблении слова «боль». Получается, «обычное использование слова служит поддержкой и для взгляда, что в отношении боли действует принцип *esse est percipi*, и для взгляда, что боль может иметь место и тогда, когда она не ощущается»<sup>1</sup>. «Но если интуиции... не составляют последовательную систему, тогда — а это именно я и утверждаю — не может быть и истинной теоретической концепции боли, и поэтому никакой компьютер или робот не может стать воплощением (*instantiations*) истинной теории боли, которая заста-

<sup>1</sup> Dennett D. Brainstorms. P. 225.

90

вила бы его чувствовать реальную боль»<sup>1</sup>. Одним словом, нам нужно вовсе отказаться от принципа *esse est percipi*, от мысли, что ощущение боли является достоверным, не подверженным коррекции состоянием, то есть фактически отказаться от *всех* «существенных» черт боли.

Где же тогда искать адекватное объяснение? Один из выходов Деннет видит в том, чтобы считать состояния боли такими же «естественными видами», какими они видятся ученым, исследующим мозг (если вообще им когда-либо удастся их увидеть), то есть состояниями, в которых нормальные причины производят все нормальные следствия. Наука будущего, возможно, покажет, действительно ли мазохисты испытывают наслаждение от боли, действительно ли общая анестезия предотвращает боль, действительно ли существуют неощущаемые болевые ощущения и т.д. Тем не менее он предупреждает, чтобы мы не обольщались относительно успехов научной стратегии. Все равно она будет основана на чем-то произвольном теоретическом решении относительно того, что есть боль. Между тем как испытывать состояние боли или не испытывать ее — это грубый факт, а не вопрос решения, зависящий от конвенций теоретика. Какой бы ни была предсказательной или элегантной теория, она не будет теорией боли в свете личных ощущений, но только теорией того, что *называется* болью. «Любое воплощение (*instantiation*) в работе любой теории боли будет уязвимым для сильных возражений, которые апеллируют к хорошо укоренившейся интуиции относительно природы боли. Все же было бы близорукостью опираться на такого рода скептические аргументы, ибо неспособность модели робота удовлетворить все наши интуитивные требования будет связана не с какой-либо некорректируемой мистикой, связанной с феноменом боли, но с некорректируемой непоследовательностью нашего обычного представления о боли»<sup>2</sup>.

Итак, что же мы имеем в итоге? С одной стороны, Деннет отмечает идею создания «чувствующего

боль компьютера» на

<sup>1</sup> Ibid. P. 227.

<sup>2</sup> Ibid. P. 228.

91

основе нашей обыденной, насквозь пропитанной противоречиями интуиции о боли. С другой стороны, он считает достаточно рискованным возлагать надежды на создание чувствующего боль компьютера с опорой только на научные теории, которые являются конвенциями теоретика. «Воплощенная» в компьютере научная теория боли вряд ли будет тождественна реальной боли. *Где* же находится «соломоново решение»? У Деннета его нет, тем не менее из его рассуждений типа «с одной стороны, с другой стороны» складывается впечатление, что его надежда связана все же с наукой.

Конечно, оговаривает он, физиологические сложности и неясности могут составить огромные и непреодолимые трудности для самых блестящих попыток теоретиков, но философские соображения не имеют к ним отношения. В том случае и тогда, когда будет разработана хорошая физиологическая субличностная теория, можно будет сконструировать робот, в котором, в принципе, ее можно было бы воплотить (instantiate). Такой прогресс в науке внесет, по всей вероятности, широкомасштабные изменения в интуитивные представления о боли. Они будут таковы, что «бремя, от которого наш робот только страдает и которое мы искусственно *называем* болью, потеряет убедительность. А тем временем, — советует Деннет, — ...интеллигентным людям стоит воздержаться от того, чтобы наносить удары такому роботу»<sup>1</sup>.

## 7. Животные и младенцы, не имеющие языка, не имеют сознания

Если согласиться с Деннетом, что признаком сознательности является способность лингвистически выражать чувства, верования, мнения и суждения, как мы должны оценивать

<sup>1</sup> Dennett D. *Brainstorms*. P. 229.

92

психику животных, не имеющих языка, а также человеческих младенцев, еще его не освоивших? Все философы, пишущие о сознании, обязательно затрагивают тему «Человек и животное», а многие из них считают, что хотя животные и младенцы не владеют языком и приобретенными паттернами культуры, им все же присущ какой-то уровень психологической жизни и рудименты сознательности. Очень часто в адрес Деннета высказывается упрек, что он проложил слишком глубокую пропасть между осознающим себя в речи человеком и лишенными языка живыми существами, включая человеческих младенцев, и одновременно смазал различие между человеком и машиной. И все потому, что для Деннета когнитивный аспект и оперирование знаково-информационной системой — более важный фактор родства, нежели физиологическое сходство. Многие пишущие о сознании авторы полагают, что ответ на вопрос «что такое сознание?» нужно искать в биологии, а не социологии, в генерировании сознания из чувствительности, присущей всем животным.

В упоминавшейся уже статье «Что это такое — быть летучей мышью?» Томас Нагель, доказывая невозможность освобождения сознательных процессов от субъективности, писал, что даже у летучих мышей следует предполагать наличие своего взгляда на мир и что они отличаются от нас только отсутствием нашего типа самосознания. Многим эта мысль показалась здравой. У Деннета иное мнение. Своим примером с мышами Нагель внес ненужную путаницу в дискуссии о сознании: у летучих мышей нет никакого сознания. Их отличие от человека состоит в том, что они не могут рассказывать истории о своем бытии, а мы можем. Летучие мыши могли бы быть сознательными на тех же самых основаниях, как и любая личность, если бы были способны генерировать тексты, на основании которых мы могли бы судить об их мире. Не только Нагель, но и многие другие философы (например, Н. Хомский, Дж. Сёрль, Дж. Фодор и К. МакГинн, Дж. Фостер и др.) не согласны с Деннетом, что в объяснении сознания мы должны двигаться от языка к биологии, а не

93

наоборот — от биологии к языку. Разгадку феномена сознания, считают они, следует искать прежде всего в биологических эволюционных механизмах и возникновении сознания человека из животной психики. Расхождение «биологицистов» с «социобиологицистом» Деннетом носит принципиальный характер, поэтому имеет смысл подробнее пояснить, как представляет себе Деннет условия появления сознания.

Для того чтобы иметь объективную картину происхождения сознания, говорит он, у нас должно быть знание о том, как в эволюции был запущен механизм «редискрипции», то есть привычка генерировать знаковые системы, формулировать с их помощью гипотезы, а также способы тестирования. Достоверного знания об этом нет ни у кого — ни у ученых, ни у философов. Однако проведенные в последние годы исследования поведения животных и вербального обучения детей позволяют сделать некоторые правдоподобные предположения. Ссылаясь на результаты Э. Кармилофф-Смита, исследовавшего освоение ребенком языка<sup>1</sup>, Деннет указывает несколько путей, на каких этот процесс может зависеть от естественного языка, даже когда напрямую и явным образом он не включает в себя использование ребенком естественного языка.

На первых шагах овладения языком звуковая речь воспринимается без понимания ее смысла. Мать говорит ребенку: «Горячо!», «Не трогай плитку!» На этом этапе ребенок не понимает значения слов «горячо», «трогай», «плитка»; для него они только звуковые события-типы, но они отзываются эхом в его памяти. В дальнейшем эти типовые события могут вызывать в его воображении типовые ситуации, и не только ту, в которой имел место некий конкретный запрет, но ситуации, в которые вовлечено определенное звуковое повторение. В процессе звукового повторения приобретает привычку *полу-понимающего само-комментария*. Под влиянием речи взрослых ребенок обретает навык сопровождать его действия

<sup>1</sup> Имеется в виду книга: Karmiloff-Smith A. A Functional Approach to Child Language. Cambridge, 1979.

94

звуковыми «комментариями», употребляя слова без понимания их значений (бессмысленный лепет ребенка). Подражая, он произносит запретительные и поощрительные слова в виде своего рода «лейблов», и постепенно эти слова приобретают семантическое значение и начинают использоваться для реального запрета или поощрения. «Однако привычка добавлять «лейблы» должна была бы появиться прежде, чем лейблы могли быть поняты, даже частично поняты»<sup>1</sup>.

Иначе говоря, звуковое подражание человеческой речи без понимания смысла Деннет считает первичным этапом обучения по сравнению с пониманием значений. По мере того, как у ребенка складываются ассоциативные связи между процессом его деятельности и звуковым сопровождением, в его памяти создаются «узелки»; слово узнается, не будучи понятым. И именно эти «якоря знакомства» являются основой, базирясь на которой «лейблы» постепенно обретают независимую идентичность внутри системы. Главным признаком независимости является использование их в *разных контекстах*. Без такой независимости смысл лейблов «невидим». То есть психика ребенка переходит от состояния простого восприятия «лейблов-репрезентаций» к состоянию понимания собственных репрезентаций. Для того чтобы для него открылась языковая перспектива, нужно, чтобы он научился отрывать «лейблы» от конкретных обстоятельств и прилагать их к разным контекстам. «Я высказываю предположение, что *слова* являются прототипами или предками *понятий*. Первые понятия, *которыми можно манипулировать*, я полагаю, являются «голосовыми» понятиями, и только понятия, которыми стало возможно манипулировать, могут становиться объектами нашего пристального изучения»<sup>2</sup>.

Свойственна ли животным способность маркировки с помощью лейблов? Известно, что многие животные разными

<sup>1</sup> Dennett Daniel C. Verbal Language. AICA conference. Oct 8, 1992. Milan // Internet: <http://ase.tufts.edu/cogstad/pubpage.htm>. P. 5.

<sup>2</sup> Ibid. P. 7.

95

способами метят свою территорию (запахами и др.). Такая маркировка предназначена, в первую очередь, для информирования других животных, она не выступает для них когнитивным



инструментом их собственной памяти, применимым для выполнения разных задач в разных контекстах. Что касается памяти, которую демонстрируют животные, делающие захоронения пищи в разных местах и маркирующие их, то это только проявление эволюционного отбора. Белкам несвойственна практика создания «лейблов» с целью облегчения разнообразных когнитивных задач и задач памяти.

Конечно, информационные системы животных очень сложны и механизмы их далеки от ясности. Остается масса вопросов без ответа. Имеется ли у собаки понятие «кошка»? «И да, и нет», — говорит Деннет. Независимо от того, насколько экстенционально близким является собачье «понятие» кошки нашему понятию, оно радикально отличается в следующем: собака не может размышлять о своем понятии, она не может задаваться вопросом, знает ли она, что такое кошка, она не может задумываться, является ли кошка животным, и т.д. В реальном мире собаки есть место для реальной кошки, но в нем нет места для понятия «кошка». Понятия находятся в человеческом мире, потому что мы располагаем особым инструментом — языком. Полярный медведь демонстрирует изоцированную компетенцию в отношении снега в его разных состояниях. Однако у него нет возможностей размышлять о снеге «вообще». Не по банальной причине, что у него для этого нет слов (естественного языка), а потому, что без естественного языка у него нет таланта вырывать понятия из их переплетенных коннекционистских гнезд и использовать в разных ситуациях. Мы можем говорить об имплицитном или процедурном знании снега полярным медведем и даже эмпирически исследовать экстенсию врожденного «снег-понятия» полярного медведя, но не понятия в нашем смысле, ибо тогда следует также предположить наличие понятия «медведь» в сознании медведя<sup>1</sup>.

<sup>1</sup> Ibid. P. 8.

96

Одним словом, Деннет считает туниковой стратегией поиска объяснения сознания только из биологии и рудиментов сознательности животных. Нередко, замечает он, биологические аргументы дезориентируют поиски. В частности, они используются для того, чтобы посеять сомнение в возможностях исследования сознания. Некоторые авторы, например лингвист Ноам Хомский, философы Джерри Фодор и Колин МакГинн, полагают, что у человека, как и у других природных существ, возможно, есть свои пределы областей исследования. Подобно тому как для пауков «закрытой» является тема рыболовства, а для львов — тема полета, почему, говорят они, нельзя предположить, что у человека тоже есть свои «закрытые» темы? Нет ли жестких природных лимитов на то, о чем могут мыслить *homo sapiens*? Почему следует думать, что в этом отношении мы отличаемся от других животных? При чем Хомский и компания, говорит Деннет, все эти вопросы представляют в виде биологических аргументов, долженствующих напомнить, чтобы мы не мнили себя ангелами без крыльев с «бесконечными и неограниченными» сознаниями. На самом деле такого рода аргументы являются «псевдобиологическими». Биология человека вплетена в паутину искусственно созданных коммуникаций, что делает ее принципиально иной системой.

«Сравнивать наши мозги с мозгами птиц или дельфинов почти бессмысленно, поскольку наши мозги в итоге соединились вместе в единую когнитивную систему, в сравнении с которой все остальное выглядит карликовым. Система соединена одной из инноваций, наводнившей наши мозги и ничьи другие: языком. Я не склонен к глупым заявлениям, что все наши мозги повязаны языком в одно гигантское сознание, обдумывающее его транснациональные мысли, но склонен считать, что каждый индивидуальный человеческий мозг, в силу его коммуникативных связей, является преемником благодеяний, полученных в результате когнитивного труда других людей и таким способом, который придает ему беспрецедентные силы. Голые животные мозги не выдерживают никакой конкуренции

97

в сравнении с тяжело вооруженными и экипированными мозгами, которые мы носим в наших головах»<sup>1</sup>.

С распространением по миру экологической идеологии (и экологических сантиментов), а также

активным обсуждением темы «права животных», позицию Деннета в отношении сознания животных порою представляют как принижение животных и реставрацию бытующей в христианстве пропасти между богоподобным человеком и остальными «тварями».

Выступая в 1995 году в Нью-Йорке на конференции «В компании с животными», где обсуждались вопросы, что значит быть животным и что значит быть человеком и как человеку следует относиться к другим живым существам, Деннет заметил, что авторами, стремящимися сблизить человека с животным, движут моральные и эмоциональные мотивы, а не философские резоны. Нормативный вопрос о социальном и морально ответственном отношении к животным и природе не следует путать с философским вопросом о сознательности животных. Под наше гуманное отношение к природе не нужно подводить ложное философское основание, а именно — признавать у животных какой-то уровень сознательности и на этом основании приписывать им моральные качества. То же самое относится к человеческим младенцам, не владеющим языком. Ответственность перед младенцами основана на моральных нормах, а не на онтологической посылке о наличии у них рудиментов сознательности.

Деннет обратил внимание на то, что некоторые люди, связывающие необходимость гуманного отношения к животным с наличием у них сознания, не приемлют мысль, что у роботов может быть признак сознательности. Асимметричное отношение к животным и роботам не имеет под собой серьезной аргументации. Например, в Массачусетском технологическом институте создан робот по имени «Cog», движения глаз и рук которого напоминают человеческие движения. Даже специалисты, знающие все о его проекте и технологии и отдающие себе

<sup>1</sup> Ibidem.

98

отчет в том, что они даже не приступили к программированию процессов высокого уровня, которые смогли бы наделить Cog сознанием, тем не менее, когда глаза робота все еще тупо следят за движениями человеческой руки, с трудом подавляя в себе чувство, что перед ними находится другой сознательный наблюдатель. Если мы признаем, что робот Cog способен своими элегантными и жизнеподобными движениями вводить нас в заблуждение, мы не должны вводить себя в заблуждение и тогда, когда мы очарованы проделками домашнего кота или пса. Общий вывод Деннета таков: «Вид информационной унификации, являющийся наиболее важным исходным условием *нашего* типа сознания, не есть что-то, что дано от рождения и что составляет часть нашего внутреннего «hard writing»; в несравненно большей мере он является артефактом нашей погруженности в культуру. Проведенное в раннем возрасте обучение прививает нам что-то вроде доброкачественной болезни, «иллюзии-пользователя», которую я назвал Картезианским Театром: иллюзии, что в нашем мозгу существует место, где происходит шоу, к которому устремляются все перцептуальные «входящие» потоки и откуда вытекают все сознательные интенции к действию и речи... Другие виды животных, в том числе новорожденные человеческие существа, просто еще не заражены иллюзией Картезианского Театра до тех пор, пока не сформировалась организация, внутри которой пользователю нужно обманываться. Данное мое предположение, несомненно, является радикальным, и многими мыслителями оно не воспринимается всерьез, им трудно даже его постичь»<sup>1</sup>.

Иначе говоря, по Деннету, для того чтобы какое-либо бытие (животное или машина) было признано сознательным, необходимо, чтобы оно обладало определенным видом информационной организации, которая снабжает его широким набором когнитивных сил (таких, как сила рефлексии и сила репрезентации). Этот вид внутренней организации не прихо-

<sup>1</sup> Dennett D.C. *Animal Consciousness: What Matters and Why* // *Social Research*. V.1. № 3. 1995. P. 703.

99

дит автоматически с так называемой «чувствительностью» (sentience) и реакцией на окружающую среду. Это не данное от рождения право млекопитающих, теплокровных или позвоночных. Это организация, которая сложилась только у одного вида животных — homo sapiens — и ни у каких других. Другие живые виды в чем-то имеют сходные нейрофизиологические организации, но

между ними и человеком различия столь велики, что апелляция к примеру других животных не имеет смысла. Как не имеет смысла приписывать признак сознательности еще не овладевшим языком человеческим младенцам. Многими философами эта идея, признает Деннет, отвергается с порога. Майкл Локвуд, защищая «картезианский материализм» и апеллируя к формуле Нагеля «что это такое, быть чем-то», заявляет: «Предположительно, сознание в этом смысле может быть найдено у всех млекопитающих и, вероятно, также у всех птиц, рептилий и амфибий»<sup>1</sup>. Вторя ему, Р. Феллоуз и Э. О'Хир утверждают, что «животные и человеческие младенцы прекрасно оперируют сознанием, не имея посредника в виде какого-либо культурно приобретенного "software"»<sup>2</sup>. Даже Пол Черчленд, наиболее близкий Деннету по мировоззрению философ, полагает, что разгадку сознания скорее всего следует искать в базисных биологических структурах, распространенных во всем животном царстве, и только на этой основе обращаться к роли языка. Авторы, ищущих ответ на вопрос «что такое сознание?» в животной психике, Деннет упрекает в отсутствии ясности и какой-либо определенности в обозначении путей, какими можно вывести из животной психики человеческое сознание. Их убеждения строятся на интуиции и предположениях, а не на фактах или доказательствах. Это и понятно. Никому не удастся найти факт, показывающий трансформацию животной чувствительности в сознание. Ни одна обезьяна,

<sup>1</sup> Loockwood M. Dennett's Mind // «Consciousness Explained». Symposium // Inquiry. V. 36. № 1/2. 1993. P. 66.

<sup>2</sup> Fellows R. and O'Hear A. Consciousness Avoided. «Consciousness Explained». Symposium // Inquiry. V. 36. № 1/2. 1993. P. 80.

100

сколько бы мы ее ни обучали, не будет применять приобретенные навыки в разных контекстах. Одним словом, Деннет настаивает на том, что если мы отстроимся от языка и культуры и будем смешивать простую реакцию животных с лингвистически оформленным человеческим чувством, тайна, окутывающая сознание, никогда не рассеется.

## Гетерофеноменология или как исследовать сознание эмпирически

Как мы уже сказали, Деннет продолжил антифеноменилистскую линию англоязычной философии и является одним из активных участников когнитивистской революции последних десятилетий. Эта революция устранила со сцены упрощенный эмпиризм неопозитивизма, заменив его более сложными схемами, принимающими во внимание многообразие когнитивных ресурсов, в том числе перспективы эволюционной истории. То есть эмпиризм не отбрасывается совсем, а переформулируется. Именно в этом свете следует оценивать его методологию. Деннет полагает, что существуют способы, позволяющие описывать сознание эмпирически, то есть таким, каково оно на самом деле, а не кажется нам. Эти способы он именуется *гетерофеноменологией*. Суть ее коротко можно представить в следующих пунктах.

— Главное отличие гетерофеноменологии от традиционной феноменологии (или *аутофеноменологии*) состоит в том, что она исходит из интерсубъективной «перспективы третьего лица» (the third-person perspective), в то время как аутофеноменология принимает субъективистскую перспективу от первого лица (first-person perspective).

— В удовлетворительном объяснении чего угодно *explanandum не должен быть частью explanans* или входить в него. Объяснить сознание невозможно, если в разряд подлежащих

101

объяснению объектов будут включаться ментальные или феноменальные объекты; они могут оказаться запутывающими исследование псевдореалиями вроде флогистона.

— К сознанию следует подходить *с нейтральной (объективной) позиции*, то есть объяснять его полностью в терминах процессов, *не являющихся сознательными*. Это значит, что исследователь не должен прибегать ни к внутренней интроспекции, ни к трансцендентным источникам и вообще не должен постулировать никакие не поддающиеся объективному анализу качества.

— Особенность гетерофеноменологии — *«метафизический минимализм»*: за скобки выводятся онтологические вопросы, относящиеся к природе сознания, и принимается методологическая

посылка, что человек есть *зомби*, а затем анализируются все релевантные факторы, обеспечивающие протекание интеллектуального процесса.

— Исследование сознания следует начинать с *когнитивной* или содержательной его стороны, а не с нейрофизиологических условий; сначала объяснять содержание (content), которое *независимо* от сознания (consciousness) и более *фундаментально*, нежели сознание, а затем исследовать сознание.

— Сознание поддается *эмпирическому исследованию*. Оно состоит в исследовании «с позиции третьего лица» того, о чем сам субъект сообщает как о своем ментальном мире, и как это проявляется в его лингвистическом и других видах поведения.

— По отношению к объектам поведения эмпиризм требует принятия интенциональной установки, рассмотрения их как имеющих определенные цели и минимальную рациональность для их достижения.

— В понимании происхождения «архитектуры» сознания, его когнитивных возможностей наилучшей стратегией является эволюционизм с его принципом естественного отбора и адаптации.

Важно иметь в виду, что, отвергая любые методы исследования сознания «изнутри», Деннет не менее критично настро-

102

ен и в отношении бытующих объективистских моделей. В октябре 1981 года в университете Миссури прошла конференция на тему «Дела сознания» (Matters of the Mind), на которой он выступил с докладом «Как исследовать человеческое сознание эмпирически». (Комментатором доклада был Ричард Рорти.) В понимании эмпирического «третьеличного» метода, говорил Деннет, до сих пор нет ясности. Эмпирические методы широко используются в экспериментальной психологии и когнитивных науках. Для получения интерсубъективной картины интеллектуальной деятельности субъектов в них применяются различные приемы — решение задач, анализ поведенческих стереотипов, исследование физических и физиологических процессов мозга и даже агрессивное вторжение в мозг. С их помощью получены интересные и проверяемые факты. Тем не менее на их основе невозможно строить теории сознания, обладающие предсказательной силой: предсказывать в рамках таких теорий — это все равно, что гадать, где и когда ударит молния.

Деннет фокусирует другие моменты. В целом можно сказать, что он переориентирует анализ с феноменальных и нейрофизиологических аспектов опыта на его содержательные, когнитивные, лингвистические и поведенческие аспекты. (В современной литературе такой подход обычно называют «когнитивизмом».) Рассмотрим более подробно, как понимается эта переориентация.

В общей форме можно сказать, что Деннет взял на вооружение схему бихевиоризма, но в отличие от скиннеровского варианта в качестве основной фактуры на «входе» берется не чувственная информация, а лингвистическое поведение — что мы говорим людям, и что они нам говорят. «Этот первый шаг влечет за собой радикальный пересмотр того, что называется данными опыта, сдвиг от акустических и физических свойств к веревкам слов»<sup>1</sup>. Издаваемые речевые шумы можно рассмат-

<sup>1</sup> Dennett D.C. How to Study Human Consciousness Empirically or Nothing Comes to Mind // Synthese. V. 53. № 2. November 1982. P. 160.

103

ривать как «содержания» или, используя модное слово, *тексты*. Анализ текстов позволяет обнаружить множество регулярностей и зависимостей. Можно обнаружить, например, что творимые субъектами тексты не являются чем-то только лингвистическим. Их содержательность может быть выражена в разных формах. Речь идет о более широком семиотическом их выражении. Например, мир «Мадам Бовари» может быть репрезентирован и в виде книжного текста, и в виде театрального действия, и видеоряда, и другими семиотическими и изобразительными средствами.

Тексты состоят не просто из произнесенных слов, а утверждений, корректировок, комментариев, то есть им присуща *интенциональная направленность*. Признание интенциональности

производимых текстов, по Деннету, вовсе не означает, как это было у Brentano, признание их необходимой связи с ментальным сознанием. Интенциональность не есть признак наличия в голове особой реальности, а показатель принятия *интенциональной установки* (intentional stance) в отношении исследуемых объектов. Данная установка имеет тот смысл, что производимые тексты *могут быть интерпретированы* как их верования и желания и что нормой последних является *рациональность*.

Следующий шаг в эмпирическом исследовании сознания — *движение за текст*. Представители разных дисциплин, имеющих дело с текстами, обычно проделывают сходные шаги. Можно обнаружить много общего между работой с текстами, проводимой исследователем сознания, работой литературоведов, занимающихся герменевтической интерпретацией текстов, и работой специалистов по искусственному интеллекту, анализирующих компьютерную информацию. Сходство в их работе состоит в возможности *отрыва от текста*. Литературный критик в принципе может двигаться за пределы созданного тем или иным автором текста и интерпретировать мир воображаемых героев — что они могли или не могли делать, — отвлекаясь как от их реальных прототипов, так и от биографии создавшего их автора. Точно так же гетерофеноменолог,

104

эмпирически исследующий сознание субъектов, может рассматривать произведенные ими высказывания (тексты) как ведущие собственную жизнь, отвлекаясь от связи этих высказываний с реальными событиями и референтами и с предполагаемым «автором» — сознанием. Главная тайна сознания, говорит Деннет, заключается в том, что *мысли сами себя мыслят*: фантазия субъекта в языковой коммуникации с фантазиями других субъектов плетет тексты о себе, о других людях, о событиях, заново переписывает их, накладывает более поздние тексты на ранние, снова редактирует и корректирует. В процессе языкового и семиотического творчества (и коммуникации) осуществляется информационное взаимодействие с действительностью. «Сознание не есть процесс, что-то из себя творящий, это состояние быть информированным — или дезинформированным — о том, что происходит в действительности»<sup>1</sup>.

Из этой констатации следует один из принципиальных выводов деннетовского *нейтрализма*: там, где мысли сами себя мыслят, не имеет значения, в каком объекте осуществляется мыслительный процесс, в ком или в чем он реализуется. Объект реализации может быть органическим мозгом или физической машиной (hardware) компьютера, или социальным институтом и организацией. Гетерофеноменолог может исследовать мысли — вымышленных обитателей текстов, — отвлекаясь от их среды обитания; ему безразлично, находятся ли они в мозгу, или в художественном произведении, или в телевизионном видеоряде, или в компьютере. Например, заложенная в компьютере информация имеет элемент самотворчества. Подобно тому, как мы можем анализировать художественное произведение с точки зрения выведения из него возможных или виртуальных событий, не предусмотренных автором, производимая компьютером информация тоже может интерпретироваться с точки зрения самотворчества и создания новой информации. Содержание этой новой информации невозможно объяснить, исходя из физических материалов, электроэнергии,

<sup>1</sup> Ibid. P. 179.

105

дизайна или исходной программы компьютера. Точно так же нас ждет неудача, если вздумаем объяснять интеллектуальные процессы человека только происходящими в мозгу физическими, химическими процессами или структурной организацией мозга.

Обычный вопрос, который задают в отношении нейтрализма, состоит в следующем: что представляет собой гетерофеноменологический мир, в котором мысли сами себя мыслят или одни тексты творят другие тексты? Этот мир, говорит Деннет, *виртуален*. Мир субъекта населен множеством воображаемых объектов, «но если мы спросим «что собой представляют эти объекты и из чего они сделаны?», мой ответ, таков: «не из чего!» Из чего сделан мистер Пиквик? Мистер Пиквик — выдуманный объект, и такими же являются объекты, описываемые, именованные,



отмеченные гетерофеноменологом»<sup>1</sup>. В отличие от биологии мозга или жесткой конструкции компьютера, у виртуальной машины сознания нет пространственно-временных характеристик. Усмотрение главного секрета сознания в том, что оно является функциональным процессом, в котором мысли сами себя мыслят, снимает, считает Деннет, давнюю оппозицию ментализма и физикализма в описании сознания. Как и при описании работы компьютерной программы, функциональные процессы сознания можно описывать на нейтральном языке, не являющемся ни физикалистским, ни менталистским.

\* \* \*

О том, как понимает Деннет виртуальность мыслей и компьютерных программ, которые сами себя мыслят, мы будем говорить в следующих главах. А сейчас попробуем суммировать рассмотренный нами выше материал.

Как говорилось в начале главы, замысел проекта Деннета состоит в опровержении дуализма с целью доказательства физического единства мира. Необходимыми условиями реали-

<sup>1</sup> Ibid. P. 175.

106

зации этого проекта является показ несостоятельности интроспективизма и феноменализма. Он убежден, что, не доказав последнее, нельзя претендовать на преодоление дуализма. Философы, думающие, что можно, не впадая в дуализм, сохранить аутофеноменологию, обманывают себя. Держась за нее, они прямиком идут не просто к дуализму, а к выведению сознания в сферу непостижимого, трансцендентного, вечной тайны, недоступной ни для науки, ни для разумного объяснения. Позицию Деннета в отношении феноменологии можно квалифицировать как элиминативизм, т.е. изъятие феноменального опыта из объяснительной схемы. В отличие от других элиминативистов он не отбрасывает сознание как химеру. Сознание можно и нужно объяснять. Однако сначала следует, во-первых, показать, чем *в действительности* является кажущаяся нам «флора и фауна феноменального сада», во-вторых, объяснить, *почему* люди верят в существование сознания и феноменальных квалиа.

Стратегия Деннета состоит в том, чтобы перевести объяснение сознания на лингвистический, когнитивный уровень. Доказать, что действительностью, скрывающейся за красочными квалиа «феноменального сада», являются *суждения* о квалиа, наши *лингвистические отчеты* на основе тех или иных реакций организма на окружающие условия, которые, в свою очередь, обусловлены природными диспозициями. Для Деннета нет различия между тем, что нам видится, и суждением о виденном: содержание сознательного опыта исчерпывается его когнитивным состоянием и вербальным отчетом. А поскольку вербальные отчеты об опыте — суждения о феноменальном опыте — являются актами лингвистического поведения, они вполне поддаются функционалистскому, объективному объяснению, то есть вполне укладываются в рамки физикалистской парадигмы.

Коли мы затронули тему физикализма, имеет смысл несколько слов сказать о направленной идентификации Деннета. Это поможет читателю лучше понять его место в философских баталиях. Сам он старается избегать определения

107

своих взглядов с помощью какого-либо «изма», считая, что всякие «лейблы» указывают на отдельные, часто формальные, стороны концепции, смазывают тонкости, не говоря уже о том, что жесткие определения предполагают схватывание «сути», что чревато аристотелевским эссенциализмом. Тем не менее очевидно, что альтернативой дуализму у него выступает физикалистский материализм.

В англоязычной философии слово «физикализм» не имеет точного определения. Оно в ходу не столько в философии физики, сколько в философии сознания. Временами оно используется как синоним «материализма», а «физическое» понимается как «материальное»<sup>1</sup>. Деннет, как и большинство физикалистски настроенных философов сознания, предпочитает не вдаваться в дебри толкования материи или физической картины мира. Он поступает проще и говорит, что в своей метафизике философам следует прислушиваться к тому, что говорит физика о мире. В его

понимании физикализм есть, во-первых, монистический принцип физического единства и непрерывности мира, исключаящий какие-либо «добавки» в виде «духов» или других эпифеноменов. То есть физике приписывается онтологический авторитет относительно того, что есть в мире, а ее законы предполагаются истинными по отношению ко всем объектам в пространстве и времени. Во-вторых, физике приписывается эпистемический авторитет: ее исследовательские нормы принимаются за стандарты получе-

<sup>1</sup> Большинство авторов все же предпочитают термин «физикализм». И вот почему. Материалистический монизм часто ассоциируется со «спинозизмом» — позицией, исходящей из *априорного* тезиса, что все на свете является материальным. Между тем современная физика (наука о материи) говорит о существовании множества физических явлений, которые не являются в строгом смысле материальными, например, о силах, волнах, полях и т.д. Называя себя физикалистами, нынешние материалисты в философии сознания, как правило, ограничиваются принятием эпистемического авторитета физики в решении проблемы «духовное—телесное».

108

ния адекватного знания о мире. Если вам угодно, то мою позицию, говорит Деннет, можно называть и «материализмом». Все, что я утверждаю, состоит в следующем: «Существует только один вид вещества, а именно *материя* — физическое вещество физики, химии и физиологии, а сознание представляет собой не что иное, как физический феномен. Короче, сознание есть мозг. Согласно материалистам, можно (в принципе) объяснить любой ментальный феномен, используя те же самые физические принципы, законы и сырые материалы, которые достаточны для объяснения радиоактивности, дрейфов континентов, фотосинтеза, репродукции, питания и роста»<sup>1</sup>. Деннет не видит большой разницы в терминах «физикализм» и «материализм». Для него только неприемлем характерный для традиционного материализма или физикализма редукционизм: объяснить содержание сознания чисто нейрофизиологическими процессами столь же невозможно, как невозможно объяснить содержание компьютерных программ физикой системного блока. Нередукционистское — нейтральное — объяснение сознания возможно только в рамках функционалистского физикализма. Эта его заявка многими оспаривается. Посмотрим, спасает ли функционализм от опасности редукционизма.

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 33.

109

### III. Сознание в рамках «Модели Множественных Набросков»

Главный вопрос состоит в следующем: как велика в мозгу интеракция между специалистами, которые определяют содержание и стиль того, что должно быть сказано, и специалистами, которые «знают слова и грамматику».

*Дэниел Деннет*

Модель Множественных Набросков Деннета дает провокативное объяснение некоторым нашим когнитивным способностям, но она не идет дальше и не говорит нам, почему вблизи этих способностей должен наличествовать сознательный опыт.

*Дэвид Чэлмерс*

#### 1. От теории тождества к функционализму «Машины Тьюринга»

В первой главе, в обзоре дискуссий по проблеме сознания, проходивших в последние полвека в англоязычной философии, мы кратко охарактеризовали основные направления наступления на картезианский дуализм. В атаках материалистов на «бастион сознания» было применено несколько стратегий:

110

трансляция психологического (ментального) языка на язык физики (неопозитивизм); перевод разговора о ментальном в разговор о семантическом и грамматическом поведении слов (логический бихевиоризм); отождествление высказываний о сознательных процессах с высказываниями о процессах мозга (теория тождества); отрицание существования реальности

сознания и ее элиминация из теоретического языка (элиминативизм); рассмотрение ментальных состояний как функциональных состояний организма (функционализм).

Деннет принял на вооружение последнюю стратегию. Он развивает радикальную версию функционализма, которую именуют по-разному: «функционализм машины Тьюринга», «AI процессуально-информационная теория», «семантический материализм», «эволюционный когнитивизм», «компьютерный физикализм» и др. В общей форме суть ее состоит в рассмотрении того, что считается ментальными состояниями человека, функциональными состояниями организма, а отношение первых к последним сравнивается с отношением «мягкой» программы компьютера (software) к его «твердой» программе (hardware). Прежде чем приступить к детальному ее рассмотрению, следует в общей форме пояснить, что понимается под функционализмом, и обрисовать контекст, определивший доминантную роль функционализма в современной философии сознания.<sup>1</sup>

Разговор о функционализме имеет смысл начать с теории тождества, опровергая которую функционализм завоевывал позиции в философии сознания. В 50-х годах Г. Фейглом, Дж.Дж. Смарт, Н. Плейсом, Д. Армстронгом и другими англоязычными философами были предложены различные варианты объяснения сознания, общим для которых был тезис о тождественности высказываний о ментальных состояниях и высказываний о некоторых состояниях или процессах мозга. Данный тезис предполагал, во-первых, перевод разговора о проблеме mind-body на языковый уровень. Во-вторых, он был прямолинейно материалистическим: главный посыл состоял в том, что каждому ментальному термину (например, терминам

111

«боль», «красное», «ревность») соответствуют референты в виде каких-то процессов мозга. Предполагалось, что описания сознательной деятельности на повседневном ментальном языке (языке фолк-психологии) и на физическом языке (например, нейрофизиологии), различаясь по типу, тождественны по своему референту. Данное тождество аналогично тождеству описаний молнии обычным наблюдателем с научным описанием ее как движения электрических зарядов. В-третьих, предполагалось, что, в то время как нейрофизиологические термины в описании процессов сознания являются каузальными и имеют законный теоретический статус в физической картине мира, ментальные термины повседневного языка (фолк-психологии) не являются таковыми: в теоретическом смысле их следует считать «номологическими бездельниками», то есть выпадающими из законосообразной картины мира.

Общий материалистический тезис теории тождества, а именно, что деятельность сознания детерминирована деятельностью мозга, мало кем оспаривался. Тезис же о том, что ментальным терминам обывденного языка соответствуют нейрофизиологические процессы, вызвал широкую критику и был признан нереалистичным. Главный вопрос, который задавался теоретикам тождества, был следующий: когда говорится о тождестве, имеется ли в виду одна сущность или речь идет о двух сущностях, одна из которых определяется через другую и в принципе сводима к ней? Если имеется в виду последнее, тогда это чистой воды редукционизм, возможность которого еще следует доказать; все предпринимавшиеся до этого попытки пока были безрезультатными. Высказывались и другие возражения: у разных биологических видов могут быть одинаковые ощущения боли, однако причины их ощущений, т.е. состояния их мозга, могут быть разными. Такое возражение высказал, например, Х. Патнэм. «Если мы сможем найти хотя бы один психологический предикат, применимый и к млекопитающим и к осьминогам (скажем, предикат "голодный"), но имеющий разные физико-химические "корреляты", то те-

112

ория состояний мозга (т.е. теория тождества — *Н. Ю.*) будет опровергнута»<sup>1</sup>.

Деннет в целом позитивно оценил желание теоретиков тождества избежать «онтологическое раздувание» мира и их твердое убеждение, что объяснение сознания не требует эпифеноменальных «добавок». В то же время он счел их тактику ведущей в ложном направлении: она напоминает ту, когда «летающие тарелки» отождествляют с болотным газом, а «русалок» с

ламантинами, то есть когда фиктивное отождествляют с реальным. Справедливо обрушившись на экстравагантный и неправдоподобный дуализм, теория тождества, замечает он, как маятник, качнулась в сторону столь же экстравагантного и неправдоподобного монизма. «Выход из этой малообещающей ситуации состоит в том, чтобы вообще отказаться от маятника. Однако это требует показа, что одно из наших первоначальных предположений не столь очевидно, как кажется на первый взгляд, а именно предположение, что, с одной стороны, существуют сознание и ментальные явления, а с другой, физические тела и явления»<sup>2</sup>. Более перспективную стратегию, способную развеять чары старых «измов» и прекратить колебания философского маятника между монизмом и дуализмом, с его точки зрения, предложил Дж. Райл в «Понятии сознания» (1949). Райл обратил внимание на то, что сознание и материя являются различными логическими категориями, и поэтому пытаться их отождествить и логически и концептуально некорректно, а главное бесполезно. Нужно искать объяснение с акцентом на логический, содержательный аспект сознания.

Самую серьезную ошибку теоретиков тождества Деннет видит в отождествлении *свойств* разных вещей — мозга и сознания, в то время как нужно отождествлять *отношения или функции*. Для пояснения он приводит такой пример. Любые

<sup>1</sup> Патнэм Х. Философия сознания. М., 1999. С. 63.

<sup>2</sup> Деннет Д.. Онтологическая проблема сознания // Аналитическая философия: становление и развитие. М., 1998. С. 362.

113

часы и любой консервный открыватель являются физической вещью, однако предположение о возможности составления предиката на ограниченном языке физики и химии, в соответствии с которым будут отобраны все открыватели банок или часы, является нереалистичным. «Какое может быть физическое свойство, благодаря которому данным дедушкиным часам, циферблатам ручных часов и солнечным часам может быть приписан предикат "показывают 10:00 утра"? В этом случае разумно только следующее утверждение: общим свойством, присущим всем открывателям для консервов, является цель или функция, независимо от их физического строения или дизайна. То же самое истинно для часов»<sup>1</sup>.

Идея о перспективности функционалистского подхода фигурирует в арсенале когнитивных средств науки и социологии не первое десятилетие. В философии сознания она выдвигалась скорее в виде спекулятивной гипотезы. Попытку подвести под нее эмпирическую основу сделал бихевиоризм с его операционалистским методом. Можно сказать, что свою родословную современный функционализм ведет именно от него. От бихевиоризма его отличает ряд существенных параметров. В бихевиористской схеме присутствует поток входящих чувственных данных (inputs), который в итоге реализуется в поведенческих реакциях (outputs), а между ними располагаются операции «черного ящика». Согласно функционализму, операции «черного ящика» следует расшифровать так: в характеристику почти любого ментального состояния обязательно входит еще и каузальное соотнесение с другими ментальными состояниями, поэтому определение ментального состояния исключительно в терминах публично наблюдаемых данных на «входе» и «выходе» (inputs и outputs) является ошибочным. Все типы ментальных состояний (ощущения, страхи, верования и т.д.) определяются их уникальными каузальными ролями в сложных взаимодействиях внутренних состояний, опосредующих чув-

<sup>1</sup> Dennett D.C. Brainstorms. Philosophical Essays on Mind and Psychology. Cambridge MA, L., 1986. P. XV.

114

ственную информацию на «входе» и проявляющихся в поведенческих реакциях «на выходе». Иначе говоря, понятие ментального состояния есть каузальное понятие. Тем самым функционалисты попытались снять одно из главных возражений против бихевиоризма, связанное с загадочностью «черного ящика», а также обвинение в редукционизме в форме сведения ментального состояния только к публично наблюдаемому поведению. Хотя функционалисты расходятся между собой в толковании онтологии каузальности, в общем и целом для них

характерен *холистический, системный* подход. Функциональная роль того или иного ментального состояния определяется всей совокупностью каузальных отношений, в которых оно находится, то есть как физическими, так и ментальными.

У функционализма есть еще одно важное преимущество. Функционалистская схема позволяет обходиться без указания на онтологию и без прямолинейного ответа на вопрос об отношении сознания и тела. Строго говоря, функционалистами могут быть и материалисты, и идеалисты, и дуалисты, ибо обработку перцептивных стимулов и их поведенческие следствия можно представить и как результат деятельности «души», и как деятельность физических сил. Преимуществом является и то, что эта схема позволяет уклониться от вопроса, принадлежит ли вся система функциональных взаимодействий конкретной человеческой личности. Вот почему многие приверженцы этой схемы называют свою позицию *метафизически нейтральной*. Правда, в случае идеализма и дуализма признание физического поведения следствием нефизических причин трудно стыкуется с законами физического мира. Большинство современных функционалистов предпочитают материализм, точнее — *физикалистскую версию функционализма*. Физикалистский функционализм отнюдь не представляет собой какую-то цельную доктрину. Существует масса его вариантов, различающихся по множеству параметров (разные толкования смысла функциональности, ментальных состояний, физических факторов, каузальности, не говоря уже о серьезных расхождениях в толковании компьютерных аналогий).

115

Хотя идея о сознании как функции высказывалась давно, по-настоящему рабочей она стала с внедрением и в психологию, и в философию компьютерных метафор и методологий, основанных на идее «Универсальной машины Тьюринга». Знакома читателя с Деннетом, мы говорили, что при подготовке им докторской диссертации (она опубликована в виде книги «Содержание и сознание», 1969), на него огромное впечатление произвел бум исследований по искусственному интеллекту. Его внимание привлекли работы Герберта Саймона и Э. Ньюэлла (1961) о симуляции компьютером мыслительных операций человека и практические результаты применения компьютерных моделей в экспериментальных работах психологов. Хилари Патнэм был одним из первых, предложивших применить функционалистский компьютерный подход к философскому пониманию сознания. В 1960 году он опубликовал статью «Сознание и машины»<sup>1</sup>. В общей форме функционализм машины Тьюринга (Turing machine functionalism) в философии сознания означает точку зрения, что каждое психологическое или ментальное состояние определяется его функциональными отношениями к получаемой чувственной информации (inputs), к результатам ее обработки во взаимодействии с другими психологическими состояниями и к проявляемым поведенческим результатам (outputs). Соответственно сознание рассматривается в виде интегрированной (холистической) системы таких состояний. Предложенное Патнэмом сравнение в дальнейшем было подвергнуто им же самим серьезной критике, тем не менее идея о возможности проведения аналогии между сознанием и компьютером была подхвачена другими философами и открыла огромную область исследований в рамках новой функционалистской парадигмы.

<sup>1</sup> Патнэм Х. Сознание и машины // Патнэм Х. Философия сознания. Впервые опубликована в: Dimensions of Mind. Ed. S. Hook. N.Y., 1960. Определение Патнэма затем вошло в словари и энциклопедии.

116

В работе «Содержание и сознание» Деннет принял «функционализм машины Тьюринга» примерно в той же трактовке, которую до него предложил Х. Патнэм.

Разумеется, что всем тем, кто задействовал компьютерные метафоры в философии сознания, не обойтись без анализа аналогий и дизаналогий естественного и искусственного интеллекта. Деннет обратил внимание прежде всего на первое. «Патнэм, — пишет он, — первый указал на замечательную и плодотворную аналогию "между логическими состояниями машины Тьюринга, с одной стороны, и ментальными состояниями человеческого бытия, с другой". Тьюринг предложил



общий способ описания организации любого компьютера или автомата в терминах упорядоченной совокупности логических состояний, которые полностью специфицированы в машинном столе их отношениями друг с другом и с input и output автомата, но чья возможная физическая реализация в hardware остается открытой<sup>1</sup>. В приведенной Патнэмом аналогии Деннет выделил ряд принципиальных моментов.

Прежде всего, применение к анализу сознания ключевых понятий: «*логические состояния машины Тьюринга*», «*отношения логических состояний*». До появления компьютеров философы сопрягали сознание либо с состоянием «души», либо с состояниями мозговых процессов. Патнэм же указал на возможность сопряжения сознания с логическими состояниями. Последние представляют собой функциональные отношения содержательных единиц, которыми оперирует сознание.

Важно отметить, что компьютерные аналогии эффективны для *когнитивистских* подходов к сознанию, т.е. когда за сознание принимается его содержательно-логическая сторона, а не феноменальная, чувственно-эмоциональная или нейрофизиологическая. Воспринятая Деннетом у Райла стратегия на редукцию сознательной деятельности к логическому, когнитивному содержанию (то, что называют «логическим бихевиоризмом») как нельзя лучше стыкуется с интеллектуализмом AI.

<sup>1</sup> Dennett D. Content and Consciousness. L., Boston, 1986. P. 102.

117

Если рассматривать сознание исключительно в когнитивном логическом аспекте, тогда различие между функционированием человеческого сознания и функционированием оснащенного софистичной программой компьютера не столь разительно. Сложнее обстоит дело, когда в понятие «сознание» включается феноменальный аспект опыта. (Об этом мы говорили во второй главе.)

Второе взятое у Патнэма важное понятие — «*реализация*» (им используются и другие сходные по смыслу термины — «инсталляция», «инкарнация» и др.). Деннет увидел многообещающую перспективу в идее Патнэма, что потенциальные каузальные взаимоотношения психологических состояний в принципе могут быть актуализированы *в разных физических структурах*. Приняв ее за ориентир, он развивает тезис о возможности реализации логического содержания в разных носителях — физических, биологических, социальных и др. «Машина Тьюринга, — пишет он, — может реализовываться непосредственно в виде электронных компонентов, или в виде "симуляции" в существующем компьютере, или в виде гидролитических волн и измерения глубины, или в виде большой комнаты, заполненной людьми, которым заданы определенные задачи»<sup>1</sup>. Принципиальный вывод Деннета состоял в том, что искусственный интеллект, имеющий дело с логическими состояниями, позволяет обойтись без обязательного сопряжения содержания сознания только с мозгом, а тем более с «душой». И вообще «материалом» реализации логических состояний могут выступать системы, где вообще нет человека как такового. Не исключено, что «материал» инопланетян совершенно иной, нежели у землян, однако для объяснения их интеллектуальных актов это не имеет значения; важна только функциональная структура внутренней деятельности, которая реализуется в их «материале». Если этот тезис принимается, в качестве такой системы можно помыслить искусственно созданную систему, могущую фигурировать как «субъект» ментальных состояний.

<sup>1</sup> Ibidem.

118

Третий сделанный Деннетом вывод состоял в том, что создание компьютеров открыло перспективу преодолеть редукционизм и обрести, наконец, *нейтральную* почву в старом споре о духовном и телесном. Главное преимущество функционализма машины Тьюринга состоит в уникальном шансе избежать устаревших «измов» и противостояния между жестким редукционизмом физикализма, с одной стороны, и менталистским дуализмом, с другой. Подобно тому как работа компьютерной программы описывается на нейтральном языке, функциональную организацию сознания тоже возможно описывать на нейтральном языке, не являющемся ни физикалистским, ни менталистским. Нейтральность языка поясняется им следующим образом:

«Для каждого ментального предиката "М" существует некоторый предикат "Ф", выразимый на некотором физически *нейтральном* языке, но спроектированном для спецификации абстрактных функций и функциональных отношений. Очевидными кандидатами на такой язык были бы системы, используемые для описания компьютеров или программ. Функциональная структура компьютерной программы может быть описана абстрактным способом и вне зависимости от какого-либо специфического описания физической "твердой" программы компьютера. Самым общим функциональным языком является система, описывающая компьютеры как "Машины Тьюринга"»<sup>1</sup>.

В функционализме Патнэма, оперирующем логическими, а не телесными состояниями, Деннет увидел реальный путь преодоления тупиков теории тождества. Двум агентам, находящимся в ментальном состоянии (например, верящим, что снег белый), нет необходимости быть физически подобными. Им достаточно находиться в «функциональном» состоянии, специфику которого можно описать в виде машин Тьюринга, т.е. признать, что оба агента находятся в некотором особом логическом состоянии. Сознания двух агентов, верящих, что снег белый, уподобляются двум разным компьютерам, имею-

<sup>1</sup> Dennett D. *Brainstorms*. P. XV—XVI.

119

щим одну и ту же программу и находящимся в том же самом «месте» в программе. Характерная для теории тождества редукция ментальных предикатов к физическим в данной позиции заменяется на редукцию ментальных предикатов к предикатам машины Тьюринга.

Деннет все же посчитал функционализм Патнэма чересчур сильной заявкой. Предположение о возможности существования некоторого особого способа описания всех агентов, верящих, что снег белый, в виде машин Тьюринга, да еще таким образом, чтобы они находились в одном и том же логическом состоянии всякий раз, когда им приписывается ментальное свойство, в лучшем случае только надежда. «На самом деле у нас нет больших оснований верить в то, что и вы и я "имеем одну и ту же программу" в каком-либо облегченном и абстрактном смысле, учитывая разницу наших природных данных и воспитания, как нет оснований верить в то, что наши мозги имеют идентичные физико-химические описания»<sup>1</sup>. Для того чтобы ослабить требования, предъявляемые ко второму ответу, Деннет предложил «функционализм признаков» (token functionalism). В качестве ответа на вопрос, что общее имеется у двух субъектов, верящих, что снег белый, им дана скорректированная формула: когда X верит, что снег белый, это значит, что этому X можно *приписать* верование, что снег белый. То есть коррекция состоит в том, что для понимания ментальных феноменов (я верю и т.д.) вводятся понятия *интенциональная система и интенциональная установка*. Если принять ментальное поведение объектов за интенциональную систему, по отношению к которой возможно применение интенциональной установки, «это позволило бы редуцировать ментальное — сферу, чьи границы в *лучшем случае* фиксируются общим признанием и имеющейся у людей интуицией, — к ясно определяемой сфере сущностей, чьи принципы организации известны, относительно формальны и систематичны»<sup>2</sup>.

<sup>1</sup> Dennett D. *Brainstorms*. P. XVI.

<sup>2</sup> Ibid. P. XVIII.

120

Деннет специально оговаривает, что концепция интенциональных систем, введенная им в классическую функциональную теорию, является неформальной и несистематичной в сравнении с концепцией Тьюринга, однако сфера, которую он хочет систематизировать, — обычные психологические состояния, описываемые на ментальном или интенциональном языке, — тоже является смутной, по меньшей мере в сравнении с ясно очерченным математическим полем прибежища функционалистской теории — сферой машины Тьюринга. При всей поверхностности аналогии между теоретическими полями машины Тьюринга и интенциональными системами, она может быть плодотворной при условии придания интенциональности иного, нетрадиционного содержания.

Обращение Деннета к понятию «интенциональность» — особенность его функционализма,

играющая важную роль в построении всей его системы. Он вводит понятие «интенциональность» совершенно в ином, нежели у Brentano смысле. У Brentano интенциональность выступала признаком ментального, у Dennett — это установка (stance), которая принимается по отношению к поведенческим объектам — живым или неживым — и признаком которой является приписывание рациональности в отношении поведения объекта и предсказания его действий. В таком смысле интенциональную установку можно приписывать и поведению животных, и поведению сложных машин, то есть интенциональность выводится за сферу онтологии человека. (Более детально о концепции интенциональности Dennett мы будем говорить в главе VI.) А раз так, считает он, снимается очень важное затруднение в объяснении сознания. «Полный успех этого проекта, — пишет он, — подкрепил бы физикализм очень скромного и недоктринального типа: все ментальные события, в конце концов, являются ни чем иным, как физическими событиями. То общее, что имеется между ментальными событиями (или между людьми, разделяющими менталистские атрибуты), эксплицируются через систему описания и предсказания, нейтральную по от-

121

ношению к физикализму, но как раз по этой причине полностью совместимую с физикализмом»<sup>1</sup>. Следует заметить, что еще в 60-х годах между Dennett и Pattem намечались серьезные расхождения не только в понимании функционализма, компьютерных аналогий и интенциональности, но и в понимании значимости эмпирических научных фактов в философском объяснении сознания. В статье «Сознание и машины» Паттэм писал: «Разнообразные вопросы и загадки, составляющие классическую психофизическую проблему, носят исключительно лингвистический и логический характер: если и есть в этой области немногочисленные эмпирические "факты", то они в равной мере подтверждают самые разные точки зрения. Я не ставлю себе цели доказать это утверждение в настоящей статье, но надеюсь сделать его более правдоподобным. В частности, я попытаюсь показать, что все эти проблемы возникают в связи с любой компьютерной системой, способной отвечать на вопросы относительно своей собственной структуры, и никак не связаны с уникальной природой человеческого субъективного опыта (если она действительно уникальна)»<sup>2</sup>.

У Dennett другое мнение на этот счет. Конечно, добытые учеными немногочисленные факты по-разному толкуются в зависимости от принятых установок. Верно и то, что для философа логические и лингвистические проблемы имеют первостепенную важность. Тем не менее полагать, что происходящее в науке не имеет отношения к философии, вряд ли верно. Работающие в содружестве со специалистами философы уже нашли для себя много принципиальных подсказок. Не исключено, что компьютерный подход является *реальным путем* к объяснению сознания.

У многих философов, говорит Dennett, существуют сомнения относительно способности функционализма выразить главное в сознании. Даже те, кто соглашаются, что картезианский

<sup>1</sup> Ibidem.

<sup>2</sup> Паттэм Х. Философия сознания. С. 23.

122

образ сознания с его интроспективной методологией и феноменализмом является философски нездравым и научно необоснованным, испытывают притягательность этого образа. Многие материалисты — и когнитивные ученые, и философы, отбросившие дуалистическую метафизику Декарта, сохраняют ее родимые пятна. Они не приемлют идею функций без свойств. Ряд видных философов (Дж. Сёрль, Д. Чэлмерс, Дж. Фодор, К. МакГинн, Дж. Левайн, Т. Нагель) настаивают на том, что функционалистское объяснение сознания не есть еще подлинное объяснение, что из него выпадает основа, создающая возможность его функционирования. Дэвид Чэлмерс, например, утверждает, что в нынешних дискуссиях «ключевым является концептуальный вопрос, состоящий в том, что объяснение функций недостаточно для объяснения опыта»<sup>1</sup>. Самую трудную задачу он видит не в показе того, *как* осуществляется опыт сознания, и не в описании его когнитивных, логических и поведенческих функций, а в ответе на вопрос, *почему* осуществление этих функций

сопровождается сознательным опытом. А Сёрль сетует на то, что при принятии функционалистской схемы из нее выпадает характерное для живых существ свойство осознанности своего бытия.

Деннет категорически не соглашается с такой оценкой «главного вопроса». Он проводит параллель между позицией Чэлмерса и близких ему по духу философов и позицией упорных виталистов, когда те говорили, что никакое физическое объяснение живой материи не схватывает главное — сущность феномена жизни — *elan vitae*. На самом деле, не устаёт повторять он, если вы представили объект холистически, показали, как функционирует вся его машина в контексте всех условий, вы объяснили все, и для агностицизма не остается пищи.

Отсутствие адекватного понимания функционализма, говорит Деннет, наблюдается не только у философов, но и у ученых. С самого возникновения когнитивных наук получил

<sup>1</sup> Цит. по: Dennett D.C. Facing Backwards on the Problem of Consciousness. Internet: <http://ase.tufts.edu/cogstud/pubpaee.htm>. P. 2.

123

распространение упрощенный взгляд на сознание. Считалось, что, подобно тому как главная функция сердца — служить насосом для перекачки крови, главная функция сознания — осуществлять контроль за органической деятельностью мозга. Отсюда делался вывод, что если компьютер, а не только органический мозг, сможет выполнять контролируемую функцию, то ему можно приписывать сознание. При этом когнитивные ученые ссылались на ученых-компьютерщиков, которые исходят из того, что любая функция, вычисляемая одной специфической компьютерной архитектурой, может быть вычисляемой и другой архитектурой. Грубо говоря, теплый, влажный органический мозг в принципе заменим охапкой силиконовых чипов. Отсюда был сделан другой ложный вывод: если все, что требуется, — это компютация, то можно игнорировать конкретные факты относительно нейробиологии и химии мозга.

Такой упрощенный взгляд, говорит Деннет, большое заблуждение. Для проектов объяснения сознания, как философских, так и в рамках когнитивных наук, важны все детали. Тем более что «история нейронауки последних лет может рассматриваться как серия триумфов для любителей деталей»<sup>1</sup>. У многих когнитивных ученых и философов, обескураженных необходимостью вникать в детали, возникло сомнение относительно самой идеи функционализма и ложное отождествление ее с общим взглядом на сознание или приблизительной схемой, не вникающей в частные факты. «Это далеко не так, — говорит Деннет. — Наоборот, основания для принятия этих новых требований являются как раз основаниями функционализма. Нейрохимия имеет значение потому — и *только* потому, — что мы открыли, что множество пронизывающих мозг разных нейромодуляторов и других химических посланий выполняют функциональные роли, которые вносят важные различия. То, что *делается* этими молекулами, оказывается

<sup>1</sup> Dennett D. Are we Explaining Consciousness Yet? // Internet. Aug. 27. 2000. P. 11.

124

важным для компютационных ролей, которые выполняют нейроны, так что все это мы должны принимать во внимание»<sup>1</sup>. Для правильного понимания функционализма важны факты, касающиеся специфической геометрии коннективистских связей, работы нейромодуляторов, временных ритмов мозга и многого другого. «Если вы оставили вне внимания *x*, или *y*, или *z*, вы не сможете объяснить, как работает сознание»<sup>2</sup>.

К сожалению, говорит Деннет, путаница относительно функционализма, имеющая место у некоторых когнитивных ученых, влияет и на философов, которые используют ее как аргумент в пользу восстановления в правах квалиа, «точки зрения от первого лица» и «субъективной реальности». Нет сомнения, говорит он, что теория сознания должна быть теорией сознания субъекта. Однако свойства, которыми наделен этот субъект, должны быть распределены в пространстве и времени и представлены в виде сложных функциональных организаций и диспозициональных характеристик. Без них «духи снова будут стучаться в дверь, требуя открыть ее для пришельцев, замаскированных под квалиа, под феноменальное или воображаемое различие между нами и зомби»<sup>3</sup>.

## 2. «Модель Множественных набросков»: децентрализация сознания

Конечно, представить сознание в виде деятельности, осуществляемой бессознательными информационно-процессуальными модулями мозга, а не как какое-то определенное находящееся в мозгу свойство, говорит Деннет, очень трудно. К тому же переход от до-сознательного уровня на уровень со-

<sup>1</sup> Ibid. P. 11—12.

<sup>2</sup> Ibid. P. 11.

<sup>3</sup> Dennett D.C. Facing Backwards on the Problem of Consciousness. Internet: <http://ase.tufts.edu/cogstud/pubpage.htm>. P. 2.

125

знательности таит в себе много загадок. Здесь можно только строить гипотезы — желательно радикальные, но в то же время правдоподобные с точки зрения имеющихся научных данных.

Для осуществления радикальных шагов нужны другие — «демократичные» — метафоры. Образ сознания в виде «Картезианского Театра» (или «стандартную» модель) он предлагает заменить принципиально иным образом. Чаще всего он называет его «Моделью множественных набросков» (The Multiple Drafts Model). Есть и другие метафоры: «Мягкая программа компьютера» (Software), «Виртуальная машина» (Virtual Machines), «Пандемониум Гомункулусов» (Pandemonium of Homunculi). В последнее время ему импонирует метафора «fame» — «известность», «слава», «знаменитость», подразумевающая абсурдный смысл известной фразы Энди Уоррела: «В будущем каждый будет знаменит в течение пятнадцати минут». Но одно дело — удачные метафоры, другое дело — теоретическое объяснение сознания в его отношении с мозгом, а именно на это претендует Деннет.

Первое, с чего следует начать строить альтернативный образ сознания, — произвести его *децентрализацию*, включающую в себя децентрализованную картину работы мозга при производстве сознания. Для этого раз и навсегда отказаться искать в мозгу авторитарную власть; здесь нет ни Короля, ни Олигарха, ни официального Куратора государственной телевизионной программы. Мозг организован более демократично и в чем-то анархично. Во-вторых, нужно признать, что сознание не есть какое-то конкретное свойство, а только функциональная организация, необходимая для выполнения информационно-коммуникативных целей. В-третьих, следует допустить, что информационно-процессуальная работа распределена в мозгу как в пространстве, так и во времени. Ошибка многих состоит в том, что деятельность сознания им представляется особым видом работы, отличающейся от работы, осуществляемой просто бессознательными информационно-процессуальными модулями в мозгу; они видят в ней творчество некой «добавки», которую, в принципе, можно было бы «субстрагировать»,

126

оставив в стороне когнитивную компоненту. Без этой «добавки», говорят они, мы будем иметь дело не с человеком, а с бессознательным зомби<sup>1</sup>.

Рассмотрим более детально, как Деннет мыслит децентрализованную работу мозга при производстве сознания. В «Сознании объясненном» он пишет: «Согласно модели множественных набросков все вариации восприятия и, конечно, все вариации мысли или ментальной активности осуществляются в мозгу параллельными, мультитрековыми процессами интерпретации и обработки чувственных inputs. Поступающая в нервную систему информация подвергается непрерывной "редакторской ревизии"»<sup>2</sup>. В соответствии с этой моделью картина происходящих в мозгу процессов предстает в следующем виде. В мозг поступает поток информации из внешнего мира или из организма, он ее перерабатывает, делая более специфицированной и дифференцированной. В процессе переработки производятся как бы черновые наброски или драфты (drafts). Сегодня в когнитивной науке, говорит Деннет, еще нет устоявшегося термина для них — носителей информации или содержания о чем-то, например, о цвете на сетине глаза, о слуховом феномене, о новизне другой информации, проходящей через мозг. Используются термины «сигнал», «содержательная фиксация», «репрезентация» и другие. Деннет предпочитает более красочные метафоры — «драфты», «бессловесные нарративы», «агенты», «демоны»,



«гомункулусы». Но суть не в терминах. Важно то, что драфты генерируются разнообразными частями мозга и в разных сочетаниях (отсюда Multiple Drafts). В итоге образуется не льющийся в одну сторону «поток сознания», а накладывающиеся друг на друга сети «бессловесных нарративов».

Реакция на получаемую информацию и последующая ее дифференциация происходят непрозрачным для сознания спо-

<sup>1</sup> Dennett D.C. *Consciousness: More like Fame than Television* // For Munich Conference volume, Feb 26, 1995. Internet: <http://ase.tufts.edu/cogsturi/pubpage.htm>. P. 1.

<sup>2</sup> Dennett D. *Consciousness Explained*. P. 111.

127

собою. Разные участки мозга, реагируя одновременно на несколько потоков информации, могут производить противоречивые драфты. Некоторые из содержательных драфтов вскоре исчезают без следа. Другие оставляют следы на последующих вербальных отчетах об опыте и в памяти: на «семантической читаемости», на эмоциональных состояниях, поведенческих склонностях. В свою очередь, они могут переписываться либо частично, либо целиком. Мультитрековые потоки по-разному несут свой специфический вклад в содержание сознания. В сотые доли миллисекунд они вносят различные добавки, осуществляют соединение, регистрацию, исправление, переписку содержания, причем все это может протекать в различном порядке. То есть в течение мгновений набрасывается что-то вроде множества сетей, которые, в свою очередь, становятся объектами непрерывного редактирования со стороны других процессов мозга. Возникающие содержания подвергаются ревизии и корректировке и в измененном виде вносят свой вклад в интерпретацию других содержаний. Важно то, что вектор этих процессов направлен в неопределенное будущее, а их эволюционный смысл — поддержка информационного контакта с информацией, идущей изнутри организма и из внешней среды с целью предугадывания ее возможных изменений и поддержания в готовности адаптационных механизмов. Обычно обращают внимание на эпистемический смысл этих процессов, на самом деле, говорит Деннет, «природа занята строительством не эпистемических машин, а машин для выживания».

В силу мультиплицированности, этот запутанный клубок содержаний — черновых набросков выглядит чем-то вроде бессловесных нарративов. В любой точке времени генерируются множество набросков фрагментов нарративов, находящихся в различных частях мозга на разных стадиях редактирования. Перепутанные сети нарративов могут прерываться на ранней или поздней стадии редактирования и производить разные эффекты. В то время как некоторые нарративы исчезают незамеченными, другие, более настойчивые, будут добиваться

128

влияния на модули внутренних состояний и внешнего поведения, а самые упорные заявляют о своем присутствии «через пресс-релиз в виде вербального поведения»<sup>1</sup>. Глобальная доступность информации в рабочем пространстве функционирования мозга и есть то, говорит Деннет, что мы субъективно испытываем как состояние осознанности или сознание<sup>2</sup>.

В соответствии с Моделью Множественных Набросков Деннет, например, следующим образом рисует картину функционирования мозга в случае *визуального восприятия*. Зрительное ощущение вызывает в коре мозга цепь реакций со все возрастающей дифференциацией и спецификацией. В разное время и в разных частях мозга принимаются различные «решения», т.е. различные части мозга каузально приведены в состояния распознавания различных черт объекта. Сначала мозг просто реагирует на поток стимулов, затем возникают последующие реакции, которые реализуются в драфтах, фиксирующих локализацию, форму, цвет, движение объекта. В итоге получается зрительное распознавание объекта. Распознающие состояния мозга передают продукты своей деятельности другим частям мозга, а те, в свою очередь, производят дальнейшее распознавание и спецификацию.

Деннет часто ссылается на работы когнитивных ученых и нейроученых, в которых приводятся данные в пользу децентрализованной картины мозга. Много интересных для себя мыслей, в частности «нарративную» идею сознания, он нашел в работе Марвина Минского «Общество

сознаний» (1985), в которой тот обосновывает тезис, что в мозгу трудится не один-единственный специалист, а «сообщество специалистов». Близкие ему взгляды он нашел в работе когнитивного психолога Б. Баарса<sup>3</sup> и других ученых.

<sup>1</sup> Dennett D.C. *Consciousness Explained*. P. 135.

<sup>2</sup> Dennett D. *Are We Explaining Consciousness Yet?* // Internet. Aug. 27, 2000. P. 1—2.

<sup>3</sup> Baars B. *A Cognitive Theory of Consciousness*. Cambridge, 1988.

129

Естественный вопрос относительно Модели Множественных Набросков состоит в следующем: почему переплетающиеся и часто противоречивые драфты не образуют хаоса и какофонии, и человек с нормальным мозгом действует разумно и по обстоятельствам? На это Деннет отвечает так: переплетающиеся драфты не образуют какофонии потому, что на их работу накладывают свой отпечаток диспозиции, точнее, два типа диспозиций. Во-первых, содержание производимых мозгом драфтов зависит от *биологически встроенных диспозиций*, истоки которых восходят к спроектированному Матерью-Природой дизайну человека. Какие-то моменты этого дизайна являются общими для всего животного мира, некоторые диспозиции включают в себя сформировавшиеся в процессе эволюции притягательность к одним вещам и неприятие других вещей. Во-вторых, деятельность человеческого сознания осуществляется на основе *приобретенных диспозиций*. Они сформировались под воздействием паттернов социума и обучения. Принципиальный тезис Деннета состоит в том, что эти социальные паттерны или микропривычки играют *решающую* роль для производства сознания. К ним относятся диспозиции производить речевые комментарии о происходящем внутри и вне человеческого бытия. Эти комментарии в значительной мере находятся под влиянием текстов, имеющих хождение в социальном окружении. Сложившиеся в культуре паттерны воздействуют не только на ход и характер «эпидемии гомункулусов», но и на эволюционную структуру мозга.

Доминантными паттернами, влияющими на все уровни мышления, по Деннету, являются паттерны языка, а еще шире — паттерны семиотических систем. Одна из его фокусных идей состоит в том, что язык — активный агент, формирующий наше бытие, в его стихии мы творим и постоянно заново *сотворяем себя* и свое сознание. Когда одна часть мозга пытается вступить в коммуникацию с другой, слова выступают катализаторами, ускоряющими передачу содержания сознания. Структуры грамматики, дисциплинируя привычки мышления, формируют диспозиции, с помощью которых мы

130

апробируем наши «базы данных». Структуры историй, приобретаемые в процессе социального обучения, руководят нами и побуждают задавать себе и другим людям релевантные обстоятельства вопросы. Разнообразные знаковые системы, используемые в организации социальной жизни, в искусстве, в науке, в идеологии определяют компетенцию мозга адекватно реагировать на окружающую среду.

Несмотря на то, что наше Я тяготеет к миражам Картезианского Театра, несмотря на иллюзии относительно собственного внутреннего мира, на практике в общем и целом мы способны к *адекватной презентации* внешнего мира. Адаптивный талант нейрофизиологических механизмов мозга и его пластичность к влиянию языка и социальных паттернов свидетельствуют о том, что все эти сложные механизмы были селективно сформированы в ходе естественного отбора и обучения в социуме. Поэтому Деннет считает вполне правдоподобным предположение, что имеющаяся у нас внутренняя модель внешней реальности, какой бы она ни была, в целом соответствует объективной реальности. Свидетельством чего является наша, в целом адекватная, поведенческая реакция на окружающую среду.

Представить себе, как эти сложные информационные взаимодействия и адаптивные механизмы может контролировать картезианский дух или отдельный нейрон, как полагают «картезианские материалисты», невозможно. В последнее время много споров, в особенности в философии физики, ведется относительно статуса «Наблюдателя». Это тоже эхо картезианства. По Деннету, бессмысленно искать пространственно-временную локализацию «Наблюдателя». «Наблюдателя»

нужно мыслить по-другому: «Когда одна порция мира (в мозгу — *Н. Ю.*) таким образом приходит к образованию клубка нарративов, эта порция мира и есть Наблюдатель. Это именно тот смысл, который может быть у наблюдателя мира, чего-то похожего на то, что что-то есть в нем»<sup>1</sup>. То есть Наблюдатель

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 137.

131

существует, однако он представляет собой не какое-то свойство самости, а «клубок нарративов», возникающих в результате информационного функционирования мозга в природной и социально-языковой среде.

### 3. Аномалии восприятия времени: «сталинская» и «оруэлловская» модели его редактирования

В последнее время, говорит Деннет, учеными зафиксированы факты, свидетельствующие о том, что порядок элементов субъективного опыта времени выглядит аномальным и даже парадоксальным в сравнении с соответствующей временной активностью мозга. Склоняющиеся к дуализму ученые (Дж. Экклз, 1977) видят в этих аномалиях свидетельство нематериальности сознания, интеракцию которого с мозгом с физической точки зрения объяснить невозможно. Не склонные к дуализму ученые (Э. Гудмен, 1978; Б. Либет, 1985) предлагают интерпретации, допускающие нарушение временного порядка причины и следствия. А совсем недавно авторитетный ученый Роджер Пенроуз в нашумевшей книге «Новое сознание императора» (1989) высказал еще более сильное предположение: материалистическое объяснение этих феноменов невозможно без революционного изменения физики и наших представлений о материи.

На самом деле, говорит Деннет, зафиксированная аномалия временных феноменов — кажущаяся. Она возникает в результате концептуальных ошибок, загоняющих исследователей в ловушку дуализма. Стандартная модель Картезианского Театра и Модель Множественных Набросков рисуют разные картины субъективного отсчета времени. Первая модель постулирует место в мозгу, «где все сходится вместе», и где распознавания во всех модальностях каким-то образом проходят регистрацию и «предстают» перед центральной самостью

132

для ее субъективного суждения. В частности, стандартная модель подразумевает, что нагруженные содержанием временные свойства событий, происходящие внутри этого привилегированного представительского медиума, детерминируют субъективный порядок. В противовес ей Модель Множественных Набросков «исходит из того, что в то время как события мозга, распознающие различные перцептуальные содержания, распределены в мозгу как в пространстве, так и во времени, и в то время как временные свойства этих различных событий являются определенными, ни одно из этих временных свойств не детерминирует субъективный порядок, поскольку не существует единого конституирующего "потока сознания", но скорее имеет место параллельный поток конфликтующих и постоянно ревизуемых содержаний»<sup>1</sup>.

Поскольку акт восприятия сразу же попадает в клубок перцептуальной памяти и других процессов мозга, отражающихся друг в друге и перекрывающихся друг друга, возникают по меньшей мере два вопроса. Можно ли вообще провести различие между восприятием событий, имевших место в действительности, и тем, что осталось в памяти от этих событий? Есть ли соответствие между: а) объективным временным порядком события; б) субъективным порядком его восприятия; в) временным порядком процессов мозга?

Деннет утверждает, что в рамках единого настоящего времени временной порядок следования сознательных состояний и событий, опосредованный состояниями и событиями мозга, не имеет нужды соответствовать воспринимаемому субъективному порядку. «Репрезентация последовательности в потоке сознания представляет собой результат интерпретативных процессов мозга и не является прямой рефлексией последовательности событий, составляющих этот процесс»<sup>2</sup>. Субъектив-

<sup>1</sup> Dennett D. and Kinsbourne M. *Time and Observer: the Where and When of Consciousness in the*

Brain. Internet: <http://ase.tufts.edu/cogstud/pubpage.htm>. P. 1.

<sup>2</sup> Ibid. P. 30.

133

ный порядок событий сознания зависит не столько от их объективного порядка, сколько от способа кодировки временных отношений в мозгу. Мозгу нет необходимости использовать объективные временные отношения для целей декодирования временных отношений<sup>1</sup>. И в этом нет никакой мистики, это эволюционное приобретение мозга. Поскольку прагматическая ценность восприятия состоит в осуществлении надежной ориентации во внешней среде, разумно предположить, что мозг стремится «метить» наши восприятия способом, который скорее отражает объективный порядок внешних событий, породивших эти восприятия, нежели их объективный временной порядок в мозгу.

В последние десятилетия в экспериментальной когнитивной психологии накопились любопытные факты, говорящие о том, что человеческий мозг ловчит с непосредственным восприятием временных событий и подделывает свою память. По поводу этой трюкаческой способности мозга в когнитивной науке сложились две объяснительные модели. Согласно одной, подделка происходит на уровне первичного акта восприятия, согласно другой — она является результатом более поздних процессов, связанных с привлечением ресурсов памяти. Деннет образно маркирует первую модель «сталинской» (имея в виду сталинские показательные процессы 30-х годов, когда вердикт осужденным выносился загодя), а вторую — «оруэлловской» (имея в виду переписку истории Министерством Истины в романе Дж. Оруэлла «1984»). Согласно «сталинской» модели, «редакторы» подвергают цензуре поступающий поток информации уже на стадии ввода информации (input), и сознанию передается уже отредактированный «текст». Согласно «оруэлловской» модели, редактирование происходит ретроспективно путем переписки в сознании полученной информации (input) в соответствии с имеющейся в памяти прошлой историей. Есть компромиссные интерпретации. Майкл Локвуд полагает, что содержание восприятия обычно включает в себя

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 64—65.

134

как сиюминутные события, так и память; то есть восприятия *окрашены* памятью, как краткосрочной, так, возможно и долгосрочной. Поэтому столкновение сиюминутного восприятия с памятью лучше рассматривать как «край» или «границу», где происходит вспышка сознания. «Редактирование» может происходить и в начале и в конце.

У Деннета позиция другая. Между этими моделями нет существенных различий в том смысле, что и «сталинская» и «оруэлловская» модели неадекватно рисуют временной порядок хода объективных событий, субъективного восприятия и протекания процессов мозга. Согласно Модели Множественных Набросков провести различие между непосредственным восприятием и его последующими обработками невозможно. Ни субъекты, рассказывающие о своих внутренних ощущениях, ни ученые, исследующие сознание извне, не располагают фактами, свидетельствующими о возможности такой дифференциации и о проведении четкой границы в мозгу между неосознанными и сознательными событиями.

Для подкрепления этого аргумента он придумал следующий пример. У меня отложилось в памяти, что я только что видел стремительно промелькнувшую мимо меня женщину в очках, с длинными волосами. На самом деле у этой женщины не было очков; в своем сознании я совместил ее образ с образом коротко остриженной женщины в очках, которую я видел раньше. «Оруэлловская» модель объясняет это тем, что у меня уже было быстротечное сознательное восприятие женщины с длинными волосами без очков; оно послужило толчком к тому, чтобы в памяти всплыл образ коротко остриженной женщины, и в итоге сформировалась память (некорректная) о длинноволосой женщине в очках. Согласно «сталинской» объяснительной модели, образ коротко остриженной женщины всплыл еще до того, как восприятие длинноволосой женщины по времени могло достичь сознание, поэтому с самого начала я увидел ее в очках<sup>1</sup>.

<sup>1</sup> Ibid. P. 117—118.

135

Этот пример показывает, считает Деннет, что не имеет смысла говорить, какая из этих моделей является корректной, поскольку у нас нет возможности судить, было ли у меня когда-либо вообще осознанное восприятие длинноволосой женщины без очков. Не существует никаких научных данных, позволяющих объективно ответить на вопрос, что в действительности и именно в данный момент испытывает осознающий себя субъект в опыте. Если рассматривать сознание в виде мультитрековых, перекрещивающихся, непрерывно саморевизируемых процессов, таких данных и не может быть. Задавать вопрос о том, что именно считать реальным, это все равно как задавать вопрос о том, каков *подлинный, канонический* текст статьи, когда автор снова и снова ее редактирует и распространяет среди знакомых ее более поздние варианты. Благодаря электронной почте и Интернету, сегодня это происходит сплошь и рядом: под влиянием советов коллег и самокритики автор часто корректирует и исправляет черновой набросок статьи (драфт) и вновь запускает его в интермедиа, желая получить отклики. Зачастую там циркулируют одновременно несколько вариантов текста одной и той же статьи. Можно, конечно, принять за канонический вариант формальный момент публикации текста и внесение его в библиографию автора, однако кто-то из читателей, возможно, посчитает первый вариант лучшим и станет работать именно с ним, и именно он будет впоследствии принят за канонический. Одним словом, публикация — это чисто произвольная акция. «Подобным же образом — и это фундаментальный вывод из Модели Множественных Набросков — если кто-то захочет установить какой-то момент в происходящих в мозгу процессах как именно момент осознания, он должен отдавать себе отчет в том, что он будет произвольным. Всегда можно условно "провести линию" в процессуальном потоке мозга, однако не существует функционального различия, которое могло бы мотивировать заявление, что все предшествующие стадии и ревизии являлись неосознанными или досознательными корректировками, а все последующие исправления содержания

136

(как они видятся в воспоминании) — как пост-опытные искажения под воздействием памяти»<sup>1</sup>. Поскольку у нас нет критерия для выбора между про-актуальным «сталинским» редактированием и ретроактивным «оруэлловским» редактированием, и постольку в мозгу не существует «финишной линии» (Картезианского Театра), глядя на которую можно точно измерить, произошло ли редактирование до или после «вспышки Сознания», различие между этими двумя теориями следует, считает Деннет, отнести к вербальному. И, конечно, не имеет значения, когда именно что-нибудь становится осознанным. «Логическая возможность ошибки в памяти открыта вне зависимости от того, насколько краток временной интервал между действительным опытом и последующим воспоминанием — это как раз то, что дает оруэлловским теориям их лицензию. ...Ошибка, закрадывающаяся в последующее верование, происходит благодаря оруэлловской памяти, а подделка является не отличимой — как извне, так и изнутри — от ошибки, которая закрадывается в первоначальный опыт благодаря сталинской иллюзии — конструкции»<sup>2</sup>.

### Поиск значений: обратная связь невербального и вербального

Самое загадочное звено в «архитектуре мозга» — переход довербальных процессов на вербальный уровень, иначе говоря — процесс стыковки биологического и социального. Это, конечно, загадка из загадок. Ее разгадывание сопровождалось множеством спекуляций и гипотез относительно того, как происходящие в глубинах мозга бессознательные процессы могут переходить на уровень мысли и словарных выражений.

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 126.

<sup>2</sup> *Ibid.* P. 318—319.

137

В дуалистической модели эту роль выполняет самость (субъект, агент, душа), выступающая генератором языка и смыслов. Самость наделяет поступающую извне и изнутри организма информацию значениями и вербальной формой: например, информацию, стимулированную тактильными ощущениями, она оформляет в понятия «горячо», «острое» и т.д. В большинстве



моделей «картезианских материалистов» фигурирует гипотеза, что где-то в мозгу должна существовать точка, где перцептуальные послания встречаются с особым нейроном — «Означивателем». Принципиально новую модель отношения между внешними характеристиками высказываний и лежащими в их основе глубинными уровнями ввел Н. Хомский. В ней обращено внимание на врожденные диспозиции к усвоению языка, однако в эмпирическом отношении она оказалась трудно доказуемой.

Современная лингвистика и когнитивные науки, занимающиеся исследованием языка, говорит Деннет, концентрируют свое внимание на слуховом восприятии языка (исследуется путь от фонологии через синтаксис к семантике и прагматике). По большей части они обходят главное — *производство речи*. «Ни одна из них — ни Ноама Хомского, ни его конкурентов или последователей — не сказала чего-либо субстанциального (правильного или ложного) относительно систем производства языка. Это все равно как если бы все теории искусства были теориями восприятия искусства, не говоря ни слова о художниках, творивших его, — как будто искусство состоит из объектов *trouves*, оцениваемых дилерами и коллекционерами»<sup>1</sup>. Причину дефицита объяснений Деннет видит в существующей неясности относительно предполагаемого предмета исследования: заниматься ли анализом высказываний и восприятий этих высказываний, речевым кодированием и декодированием, или «языком мысли», или глубинными структурами. Лингвисты и когнитивные ученые могут позволить себе отложить вопросы производства речи на потом и заняться периферийными вопро-

<sup>1</sup> Ibid. P. 231.

138

сами, но философы не имеют права уходить от главного и не высказывать по этому поводу гипотезы.

Поскольку никому еще не удалось представить какое-либо ясное описание инициатора речевого процесса, итогом которого является полнокровное высказывание, говорит Деннет, трудно даже представить, с чего нужно начинать строить теорию. Трудно, но не невозможно. Об этом свидетельствуют появившиеся в последнее время интересные заявки на такие теории, обзор которых сделан датским психолингвистом Пимом Левельтом<sup>1</sup>.

Исходя из альтернативной модели — модели Пандемониума и тех исследований в психолингвистике и когнитивной психологии, которые ему представляются наиболее значительными, Деннет высказывает следующую гипотезу. Когда мы имеем дело с функционирующим мозгом, «главный вопрос состоит в следующем: как здесь велика интеракция между специалистами ("драфтами", "демонами"— *Н. Ю.*), которые определяют содержание и стиль того, что должно быть сказано, и специалистами, которые "знают слова и грамматику"»?<sup>2</sup> Конечно, прояснение этой интеракции зависит в первую очередь от прогресса в эмпирических исследованиях мозга. Тем не менее уже сейчас есть основания сказать, что бытующая «бюрократическая» модель, согласно которой содержательные драфты передаются по эстафете некоему «Центральному означивателю», наделяющему их смыслом и вербальной формой, является ложной. Хотя ясности нет и здесь. «Проблема с альтернативой Пандемониума состоит в том, что мы должны найти способ, каким источники содержания могут *влиять* или ограничивать творческие энергии слов-демонов *без диктанта* по отношению к ним»<sup>3</sup>.

Наиболее перспективным Деннет считает подход, принимающий гипотезу о взаимодействии драфтов, несущих содер-

<sup>1</sup> Levelt P. *Speaking*. 1989.

<sup>2</sup> Dennett D. *Consciousness Explained*. P. 240.

<sup>3</sup> Ibid. P. 238.

139

жательные послания, и вербальных драфтов. «А что если в параллельном процессе слова-демоны выступают вопрошателями/ограничителями, а несущие содержание демоны являются ответчиками/судьями? Хорошо развитые и натренированные коммуникативные интенции — значения — могли бы возникать из квазиэволюционного процесса дизайна речевого акта, который

включает коллаборацию, частично сериальную, а частично параллельную, различных подсистем, ни одна из которых сама по себе не способна осуществлять — или предписывать — речевой акт»<sup>1</sup>. Другими словами, гипотеза Деннета такова: связь невербального с вербальным осуществляется не на основе диктата первого или диктата второго, а на основе *принципа оппортунизма, параллелизма, взаимной коррекции и обратной связи*. Содержательные драфты влияют на выбор слов, но и слова, в свою очередь, влияют на характер несомого драфтами содержания. Поэтому жесткую демаркацию между невербальными и вербальными процессами мозга провести невозможно. «Существует не один, а множество смещающихся источников значения, оппортунистически развитых в процессе поиска адекватных слов. Вместо определенного содержания, находящегося в конкретном функциональном месте, ждущего быть англоязычным с помощью субрутин, имеет место все-еще-неполно-определенный-набор сознаний, распределенных по всему мозгу и ограничивающих процесс композиции, который в ходе времени может обрести обратную связь для проведения ревизии или приспособления, а далее для определения экспрессивной задачи, которая в первую очередь запускает в движение процесс композиции»<sup>2</sup>.

В начале этой главы мы уже говорили о том, что, критикуя теорию тождества, Деннет отверг идею соответствия процессов сознания физическим процессам мозга. Он не исключает, что ряду наших мыслей и других психологических процессов

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 239.

<sup>2</sup> *Ibid.* P. 241.

140

вообще нет физиологического соответствия. Ибо большая часть из того, что мы говорим, думаем и делаем, внутренне зависима от нашего выхватывания из массива культуры фраз, мыслей и их логических связей. Задействованность языковой компоненты делает невозможным ни сведение содержания сознания к нейрофизиологическим процессам в мозгу, ни его выведение. Это значит, что между мыслимым (когнитивным) содержанием и реально существующим миром нет изоморфизма, их логические структуры не совпадают. Правда, Деннет еще не готов полностью отказаться от психофизического параллелизма, но разговор о нем можно вести только на макроуровне эволюционного развития. Ибо согласно его метафизике мир является физическим, и все, что в нем возникает, не выходит за его рамки. Сознание — это способ функционирования физического, способ его организации, который в общем и целом соответствует алгоритмам эволюционных процессов. Итак, смысл деннетовской Модели Множественных Набросков — нарисовать децентрализованную картину работы сознания. Показать, что в нашей голове трудится не один имеющий жесткую бюрократическую систему гомункулус, а множество, и каждый сознательный акт представляет собой как бы эпидемический взрыв (пандемониум) гомункулусов, выполняющих разнообразные функции и говорящих на разных языках. Центра или главной точки, где сплетаются воедино разнообразные потоки драфтов, где они «регистрируются» и где принимается «окончательное решение», не существует. «Не существует единого, определенного "потока сознания", потому что не существует Главного штаба, нет Картезианского театра, где "все это соединяется вместе", чтобы быть представленным вниманию Центрального Означивателя»<sup>1</sup>. Поэтому и картезианские дуалисты, занятые поисками ментального «Наблюдателя», и картезианские материалисты, ищущие в мозгу «Главный нейрон», стягивающий к себе каузальные потоки информации, выстраивают свою стратегию на псевдореалиях.

<sup>1</sup> *Ibid.* P. 257.

141

Из сказанного выше читатель может судить, что Деннет не ограничивается типичными для философов-натуралистов заявлениями и стремится нарисовать более или менее конкретную картину работы «Машины Сознания». Характер высказанных им гипотез можно было бы определить старым термином «диалектика».

И все же нарисованная им картина в соответствии с Моделью Множественных Набросков далека от ясности. Обозначим только главные недоумения. Первый вопрос, конечно, об *онтологическом*

*статусе сознания.* Существует ли сознание? На этот вопрос Деннет отвечает утвердительно. Элиминативист в отношении феноменологии и квалиа, он не является элиминативистом в отношении сознания. Однако ставить вопрос об особой «онтологии сознания» не имеет смысла: сознание не является ни ментальным, ни физическим, оно нейтрально, это только функция информационно-процессуальных систем мозга, а не какое-то отдельно взятое свойство. Онтологически сознание есть «Ничто» (или «Информационная диаспора»). Роль этого «Ничто» быть *посредником* между машиной мозга и внешним миром, в существование которого люди верят и рассказывают о нем на языке феноменологии. В таком же духе он отвечает и на второй вопрос: «Где все эти мультитрековые активности мозга воссоединятся вместе?» Краткий ответ Деннета таков: «Нигде» (Nowhere)<sup>1</sup>. Цель Модели Множественных Набросков состоит в том, чтобы доказать возможность уровня описания среднего (нейтрального) пути между нейрофизиологической машиной мозга и всего того, что на языке фолк-психологии мы именуем ментальным содержанием (нашими верованиями, суждениями и другими интенциональными продуктами).

Распространено мнение, говорит Деннет, что перескочить с объективного уровня физиологии на субъективный уровень феноменологии невозможно (Т. Нагель, 1974, 1986.). Совсем недавно Колин МакГинн заявил, что у сознания скорее всего

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 135.

142

имеется «скрытая структура», находящаяся вне физиологии и феноменологии. Благодаря ей осуществляется соединение этих уровней, однако природу этой структуры вряд ли можно раскрыть. «Как мне представляется, — говорит он, — этот вид скрытой структуры не находится ни на одном из тех уровней, который обозначил Нагель: предположительно он располагается где-то между ними. Не являясь ни феноменологическим, ни физическим, этот опосредующий уровень вряд ли (по определению) сконструирован по модели той или другой стороны этого раскола (между физиологией и феноменологией — *Н. Ю.*) и в силу этого лишен способности достигать другую сторону. Для его характеристики потребуются радикальная концептуальная инновация (которая, на чем я настаиваю, вероятнее всего, находится вне наших возможностей)<sup>1</sup>.

Деннет не разделяет пессимистическое мнение МакГинна о том, что медиум между физиологией и феноменологией — *terra incognita*. Он говорит: «Мы осуществили это! Мы вообразили, как мозг производит сознательный опыт»<sup>2</sup>. Опосредующий уровень, прояснение которого МакГинну кажется «вне наших возможностей», — это «Software» или «виртуально машинный» уровень описания. Не являясь открыто физиологическим и механическим, он способен тем не менее, с одной стороны, обеспечить необходимые мосты в машине мозга, с другой стороны, не являясь открыто феноменологическим, он в то же время способен обеспечить необходимые мосты к миру содержания, миру гетерофеноменологии. «Понятия компьютерной науки дают подпорки для необходимого нам воображения, если мы хотим продраться через *terra incognita* между феноменологией, какой она представляется нам по нашей «интроспекции», и нашими мозгами, о деятельности которых рассказывает наука. Рассуждение о наших мозгах как информационно-процессуальных системах может постепенно рассе-

<sup>1</sup> McGinn C. *The Problem of Consciousness*. Oxford, 1991. P. 102—103. Цит. по: Dennett D.C. *Consciousness Explained*. P. 433—434.

<sup>2</sup> Dennett D.C. *Consciousness Explained*. P. 434.

143

ять туман и поможет увидеть путь сквозь этот великий провал, показывая, как же это может быть так, что наш мозг продуцирует все феномены»<sup>1</sup>.

Можно сказать, что в отличие от скептиков, сомневающихся в возможности преодоления пропасти между физиологией и феноменологией, Деннет настроен оптимистически. Он убежден, что с помощью предлагаемой им «Модели Множественных Набросков» медиум, который многие относят к *terra incognita*, вполне поддается объяснению, это не есть неведомая науке субстанция, это функция и способ организации эволюционирующей материи.

Принципиально важным вопросом для всякой концепции, претендующей на радикально новое объяснение феномена сознания, является вопрос о характере теоретического языка его описания. На этот вопрос у Деннета есть вполне определенный ответ: необходимый нам уровень описания и объяснения мультитрековой функциональной организации сознания аналогичен (хотя и не идентичен) тому, который называют «software level» в описании компьютеров. Главная задача исследователя состоит в том, чтобы понять, как человеческое сознание реализуется в операциях *виртуальной машины*, созданной культурой в мозгу. Если в ранних работах у него еще достаточно «слабые» формулировки, в последних он прибегает к «сильным»: определенные типы сложных компьютерных машин могли бы быть «виртуальными машинами» и обладать сознанием. Речь идет прежде всего о методологической значимости предположения о возможности таких машин: приняв его за правдоподобное, можно преодолеть пропасть между физиологией и феноменологией, которая многим кажется непреодолимой.

Деннет с удовлетворением констатирует, что в последнее время в когнитивных науках складывается более конструктивная и перспективная тенденция к преодолению магии картезианского образа. Среди многих ученых, занимающихся со-

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 433.

144

знанием, складывается что-то вроде консенсуса относительно правдоподобности модели сознания, изображающей его в виде «*глобального нейронного рабочего пространства*» (global neuronal workspace model). В целом это движение, считает Деннет, идет в правильном направлении. Сознание рассматривается не как отдельное свойство и не как медиум неизвестной природы, а как глобальное функционирование нейросетей мозга. Но и здесь, считает он, не обходится без путаницы. Некоторые видят в «глобальном нейронном рабочем пространстве» *причину* сознания, как если бы сознание еще нечто другое, дополнительное к этому пространству условие. На самом деле «Сознание представляет собой не привилегированный медиум репрезентации или добавочное свойство, имеющееся у некоторых его состояний; оно само есть именно взаимная доступность, дающая некоторым информационным состояниям силы, возникающие с осознанием субъектом этой информации»<sup>1</sup>.

Несмотря на увеличивающееся количество экспериментальных данных, опровергающих центристский образ сознания, говорит Деннет, этот образ продолжает скрыто действовать на умы. Например, в исследованиях нейрочелюв встречаются неосторожные формулировки вроде «проблемы связующих креплений» (the binding problem), уже постановка которой предполагает, что в мозгу следует искать некое единое представительское место, где проходят регистрацию результаты всех проведенных в нем различений. Другие, избегая отдающей картезианством идеи «нейрона-медиума», используют понятие «финишная линия», пересекая которую бессознательный поток информации вдруг преобразуется в сознательный. Ее местоположение никто не знает, а исследующие зрительное восприятие ученые затрудняются сказать, что же можно посчитать за «финишную линию», когда, например, мы видим красный цвет: проходит ли она через глазное яблоко, сетчатку или где-то в глубинах мозга, где запрятаны подлинники «глаза сознания».

<sup>1</sup> Dennett D.C. *Are We Explaining Consciousness Yet?* P. 1.

145

ния». Некоторые когнитивные ученые продолжают оперировать понятием «поток сознания», изображая сознание в виде непрерывной последовательности содержательных единиц, некоторые из которых неожиданно взрываются сознательными актами и в таком виде входят в память.

Ученым, как и философам, свойственно настроение ожидания. Некоторые из них полагают, что объяснение сознания если и возможно, то только с открытием принципиально новых фактов в нейронауках и в других серьезных областях знания. На науку будущего, говорит Деннет, уповают те, кто остается во власти идеи, что «сознание ... есть разновидность полумистического свойства, об открытии существования которого сможет оповестить только тиканье датчиков фенонометра или известнометра»<sup>1</sup>. Модель Множественных Набросков, убежден он, дает шанс ученым освободиться от стереотипов и увидеть новые задачи и новые перспективы их решения. Эта

модель обладает рядом преимуществ по сравнению с другими моделями: она эмпирична, в ней нет места для «чуда», с ее помощью можно эффективно справиться с такими загадками, которые ставят в тупик ученых. Например, загадками временных аномалий, несовпадения субъективного и объективного времени, производства языка, «феномена Phi» (The Phi Phenomen) и др.

В этой главе мы уже затронули важный пласт деннетовской философии — компьютерную идеологию. Она будет предметом рассмотрения в следующей главе.

<sup>1</sup> Dennett D.C. *Consciousness: More Like Fame than Television*. P. 12.

## IV. «Машина Тьюринга», «Машина Дарвина» и виртуальная «Машина Сознания»

Дарвиновская парадигма содержит все, чтобы принять и творчество человека, и AI в пространство своего дизайна, ибо ее принцип «Исследование и Развитие» (R&D) свойствен и машине.

*Дэниел Деннет*

Социальный институт языка не имеет никакого отношения к генезису сознания.

*Пол Черчленд*

### 1. Компьютерные метафоры и философия

Если сознание является не Центральным Наблюдателем, а функционированием информационных процессов, а феноменальный мир ощущений и чувств есть не что иное, как кажимости, возникает вопрос, может ли иммунный к «миражам феноменологического сада», но подражающий человеческому когнитивному поведению робот (или зомби) считаться таким же осознающим бытием, каким является человек? Ответ Деннета: «да» и только «да». Главным признаком сознания является как раз то общее, что имеется у нас с компьютерами, — функционирование когнитивно-информационных процессов. Если человеческое сознание представить в виде некоей вир-

147

туальной машины (или software), тогда любой сравнимой с ней *виртуальной машине* можно приписывать сознание. Определенные типы сложных компьютеров, называемых Искусственным Интеллектом, могли бы иметь типы виртуальных механизмов и считаться сознательными.

Такого рода заявления Деннета не следует понимать в смысле утверждения, что сегодняшние компьютеры обладают полнотой человеческих качеств. Мы поймем его превратно, если будем принимать за примитивного компьютерофила, верящего, что раскрытие тайны сознания целиком зависит от дальнейшего прогресса компьютерной техники. В феномене искусственного интеллекта его привлекает прежде всего философский аспект. По большому счету его интересует извечный метафизический вопрос о единстве мира: как возможно доказать непрерывность физического, биологического, интеллектуального (культуры)? Наиболее приемлемой перспективой он считает следующую: выявить сходство алгоритмов «Машины Дарвина», «Машины Тьюринга» и «Машины Сознания». Если это будет показано, тогда не только фантом *res cogitans* должен покинуть философскую сцену, но и «Великий раскол» сознания на нейрофизиологию и феноменологию, кажущийся многим непреодолимым, исчезнет. Эта перспектива связана у него с изначальной посылкой: не только эволюция природы обладает способностью в одном существе соединить тело и мысли, но и человек, будучи продолжением природы, творя артефакты, соединяет природу и культуру и не создает онтологических расколов.

В предыдущих главах мы уже касались взглядов Деннета на компьютерные дела. В настоящей главе мы постараемся более подробно осветить особенность его стратегии — симбиоз тьюрингианства и дарвинизма.

Будут рассмотрены аналогии и дизанalogии естественного и искусственного интеллекта, гипотеза об эволюционном «размещении» сознания в мозгу, роль языка и «меме-вирусов» в производстве «виртуальной машины» сознания, а также понимание принципа эволюционизма (алгоритма «Машины



148

Дарвина»). Еще раз хотелось бы предупредить читателя, что в обсуждении этих тем Деннет задействует самую различную информацию, в том числе специальное знание. Он ссылается на работы многих ученых, принимает или отвергает их выводы. Поскольку мы не являемся специалистами в научных областях, многое из этого материала, безусловно философски важного, мы опускаем. В его рассуждениях, повторяем, нас интересует метафизический вектор. Учитывая, что наш рассказ о философии Деннета по жанру является историко-философским, предполагающим освещение контекста, мы начнем его с краткой истории внедрения компьютерной идеологии в философию.

Классиками, положившими начало компьютерным разработкам, и наиболее известными изобретателями компьютера были британский математик Алан Тьюринг (1912—1954) и венгерско-американский математик и физик Джон фон Нейман (1903—1957). У Тьюринга была большая практика в создании и построении специальных электронных кодирующих и декодирующих аппаратов, которые помогли союзникам в перехвате немецкой информации во время Второй мировой войны, однако именно чисто абстрактные работы привели его к созданию теоретического компьютера, получившего название «Универсальная машина Тьюринга». Именно от него ведется отсчет Компьютерного века. В статье «Вычислительные машины и интеллект» (1950) он высказался в пользу логической возможности машинного интеллекта; каждая функция, вычисляемая естественным образом живым математиком, может быть вычислена с помощью модели «Универсальной машины Тьюринга». Эта модель явилась первым уточнением понятия алгоритма. Будучи мысленным экспериментом, а не инженерным проектом, она фактически была философской. Он же предложил критерий оценки интеллектуальности машины, состоящий в том, чтобы по системе получаемых ответов спрашивающий не мог бы с уверенностью судить, даются ли они человеком или машиной. В дальнейшем этот критерий получил название «Тест Тьюринга». Фон Нейман обратил внима-

149

ние на прикладной аспект абстракции Тьюринга и сделал ее достаточно конкретной для создания практически работающего компьютера. Абстрактный дизайн, известный как «архитектура фон Неймана», сегодня присутствует почти в любом компьютере — от гигантского до домашнего PC. В 60-е годы, когда компьютеры были еще в новинку и об их интеллектуальных возможностях известно было меньше, нежели о человеческом мышлении, для сенситивных к научной новизне философов и психологов фокусным был вопрос: «Могут ли компьютеры мыслить?» По мере дальнейших разработок искусственного интеллекта (AI) и достигнутых впечатляющих результатов этот вопрос разбил на огромное количество специальных вопросов. Было получено достаточно точное знание о том, что могут делать компьютеры, как они делают это, чего они пока делать не могут, но в принципе в состоянии делать, а чего они вообще делать не смогут. Во всяком случае, добытое здесь знание об интеллектуальных операциях оказалось гораздо более точным, нежели то, которое было накоплено во всех проводившихся до того исследованиях сознания — психологических или философских. По мере увеличения объема данных о способностях компьютера, когда-то широко дебатировавшийся вопрос «могут ли компьютеры мыслить?» изменил вектор на противоположное направление и обрел новые смыслы: «Является ли человеческое сознание вычислительной машиной?», «Как возможен феномен искусственного интеллекта в природе?»

Исследования по искусственному интеллекту оказали огромное влияние на самые различные дисциплины, включая психологию. Одно из ее главных изменений состояло в допущении, что объяснение мыслительных процессов человека лучше всего достигается с помощью компьютерных понятий. Компутационная гипотеза предполагает утвердительный ответ на вопрос «является ли человеческое сознание коммутационной машиной?», хотя и не в смысле метафизического утверждения, что сущностью сознания является коммутация, а в смысле применения принятых в компьютерных науках методоло-

150

гий. Компутационная методология стала считаться более эффективной в объяснении операций интеллекта прежде всего потому, что лучше поддается экспериментальной проверке. Психология, принявшая на вооружение язык и методологию компьютерных наук, превратилась в *когнитивную психологию*, стала областью когнитивных наук, исследующих мыслительные процессы относительно точными методами.

Новые веяния в компьютерных науках и когнитивной психологии сказались и на философии, в особенности на аналитической, всегда стремившейся к максимально точному анализу разумного рассуждения. В смысле техники анализа компьютеризация явилась следующим шагом на пути к точности и проверяемости, сравнимым с введением в философию логической формализации. Конечно, среди философов реакция на компьютерные метафоры не была однозначной; кто-то посчитал все аналогии естественного и искусственного интеллекта пустой тратой времени, кто-то занял осторожную тактику выжидания, а кто-то увидел в приходе компьютерного языка и методологии долгожданный ключ к тайне человеческого интеллекта и разрешению извечных философских споров.

В дискуссиях философов было задействовано множество разнообразных и в то же время взаимосвязанных вопросов. Что может, а что не может делать компьютер в сравнении с человеком? Правомерна ли аналогия мозга с hardware компьютера, а сознания — с software? Каков оптимальный язык описания деятельности сознания: феноменологии, нейрофизиологии или компьютерный язык? Какие философские выводы следуют из аналогий и дизанalogий естественного и искусственного интеллекта? Можно ли сравнивать информационные приращения компьютера с смысловой деятельностью сознания? Способен ли компьютер к наполнению синтаксиса семантикой и подлинному пониманию? Как вообще стал возможен в природе феномен AI? По большому счету все эти вопросы сводятся к главному: что такое человек в системе природы и артефактов?

151

Постановка последних двух вопросов, несомненно, имеющих широкий мировоззренческий характер, была стимулирована не только появлением феномена AI, но и новыми толкованиями эволюционного принципа. В том же 1969 году, когда была опубликована книга Деннета «Содержание и сознание», вышла в свет книга Герберта Саймона «Науки об искусственном»<sup>1</sup>, в которой представлен новый взгляд на понимание естественного и искусственного. В ней автор дал предельно широкое толкование адапционизму, распространив эволюционный принцип на артефакты или искусственные системы. Артефакты, конечно, являются результатом творчества человека, тем не менее, по его словам, они могут быть охарактеризованы в терминах функций, целей, адаптаций. Более того, он утверждал, что организмы можно рассматривать как искусственные системы, в том смысле, что, будучи адаптивными механизмами, они эволюционировали благодаря силам органической эволюции. И организации в сфере бизнеса или машинного производства, и функциональные организации биологических организмов, с его точки зрения, являются примерами адаптивных систем. Поэтому он предложил рассматривать науку об искусственном — как обобщающую науку о дизайне. Или науку дизайна. Соответственно когнитивную науку он включил в «науку дизайна», предметом которой является исследование дизайна когнитивных артефактов, созданных человеком для адаптации к среде.

Весь этот идейный контекст дискуссий о естественном и искусственном интеллекте, об универсальности или ограниченности принципа адапционизма повлиял на философию сознания, осложнил и сделал более плюралистичной ситуацию в ней.

Мы уже говорили, что в философии конфликтующие методологии сложились практически относительно всего, исследовать ли сознание «изнутри» или «извне», «снизу» или «сверху». Позиции по отношению к AI — к его имеющейся или потен-

<sup>1</sup> Саймон Г. Науки об искусственном. М., 1972. Издание на английском языке. 1969.

152

циональной объяснительной силе — зависят от принятой парадигмы: социолингвистической, биологической или социобиологической. Социологи считают, что сознание следует

объяснять «сверху» — культурой и языком; биологицы полагают, что «снизу», из присущих животному миру свойств психики, а язык и культуру относить к вторичным наслоениям; социобиологи предлагают соединить первые два подхода, применив новую методологию, позволяющую объяснять сознание одновременно и «снизу» и «сверху».

Деннет выбрал для себя последний вариант. Конечно, работать в социобиологической парадигме сложнее, ибо предметом является динамика взаимодействия природных и культурных факторов, о которой еще мало что известно. (Этим объясняется его внимание к новейшим научным исследованиям, могущим пролить свет на неясности.) С этим связан и сложный ход его рассуждений, и многоуровневые построения, и соединение тьюрингианства с дарвинизмом. Свою задачу он видит в том, чтобы нарисовать картину сознания, во-первых, в процессе и динамике, во-вторых, в дуальной перспективе биологической и культурной эволюции.

### Виртуальная «Машина Сознания»: ее «размещение» в органике мозга

Не будет преувеличением, если сказать, что анализ человеческого сознания в сопоставлении с компьютером играет одну из ключевых ролей в философии Деннета и является важнейшим источником его новаций. В понятиях компьютерной науки он увидел не только долгожданное средство для разгадки тайн, над которыми веками ломали головы философы, но и средство для раскрепощения философской фантазии от устаревших клише и одновременно стимул для полета воображения и интересных спекуляций. «При создании компьютерной модели или алгоритма импрессионизм неуместен. В то же время

153

у импрессионизма есть обширное пространство, когда дело доходит до интерпретации этой модели. Компьютеры заставляют прояснять вещи, которые нам важно представлять ясно. На деле AI является новым и лучшим средством в работе определенного типа философии»<sup>1</sup>.

Деннет относится к тем философам, которых Джон Сёрль назвал сторонниками «сильного AI». Если сторонники «слабого AI», к числу которых Сёрль относит и себя, видят в компьютере инструментальную ценность — его способность осуществлять формальную когнитивную работу, проверять гипотезы, строгость формулировок и т.п., сторонники «сильного AI» не ограничиваются инструментализмом; они полагают, что о соответствующе запрограммированном компьютере можно говорить как о некоем *сознании* в том смысле, что он способен иметь не только формальные когнитивные состояния, но и такие состояния, как понимание, саморефлексия и творчество. Мы будем говорить на эту тему в разделе «Деннет versus Сёрль».

В данной главе мы остановимся более подробно на понимании им механизма взаимодействия природной (биологической) структуры мозга и содержательной составляющей человеческого сознания или, в его терминах, взаимодействия software сознания с органической hardware мозга.

Один из принципиальных тезисов Деннета состоит в следующем: «Осознающие себя человеческие сознания являются более или менее сериальными виртуальными машинами, имплементированными — непроизводительно — на параллельном hardware (мозга — *Н.Ю.*), который нам предоставила эволюция»<sup>2</sup>. Этот тезис требует пояснения. Поразившие воображение современников машины фон Неймана, говорит он, были названы «гигантскими электронными мозгами». На самом деле их правильнее было бы назвать «гигантскими элек-

<sup>1</sup> Из интервью Д. Деннета. 07. 02. 2000. Интернет: <http://www.theatlantic.com/unbound/digitcult/dennett.htm> . P. 1.

<sup>2</sup> Dennett D.C. Consciousness Explained. L., 1991. P. 218.

154

тронными сознаниями» (minds), упрощенными электронными имитациями того, что Уильям Джемс называл «поток сознания». Исторический факт состоит в том, что при создании своей машины фон Нейман апеллировал к интроспекции, к мыслительным шагам, которые, с его точки зрения, обычно предпринимаются разумным человеком для получения искомого результата. Поэтому архитектура его машины тоже была сериальной — пошаговой. На самом деле, о чем свидетельствует наука, в отличие от льющегося в одну сторону «потока сознания» Джемса,

архитектура мозга является параллельной, в ней одновременно задействованы миллионы каналов и операций. Мозг человека больше похож на параллельный, нежели сериальный процессор. (Они составляют основу большинства использующихся сегодня компьютеров.) Различие между ними состоит в том, что когда сериальные процессоры выполняют в единицу времени одну задачу, параллельные процессоры — множество разнообразных задач.

При всей заманчивости компьютерных аналогий, говорит Деннет, между параллельной архитектурой мозга и (стандартной) сериальной архитектурой компьютера существует огромное различие. Критики обычно используют этот факт в качестве одного из самых фундаментальных возражений против аналогии сознания с искусственным интеллектом; в отличие от мозга последний, говорят они, почти всегда работает как машина фон Неймана, то есть сериально. Если исходить из того, что наш мозг является тысячеканальным параллельным процессором, описывать его шаги проблематично, поскольку невозможно иметь интроспективный доступ к процессам, одновременно идущим по тысячам каналов, даже если они происходят в собственном мозгу. Иначе говоря, хотя архитектура мозга имеет параллельный характер, интроспективный доступ к происходящим сознательным событиям в мозгу осуществляется в формате последовательности. Поэтому и для когнитивных ученых, и для философов, говорит Деннет, главный вопрос состоит в следующем: каким

155

образом фон-неймановский сериальный феномен мог возникнуть в параллельной и многоходовой разноголосице мозга? Как получилось, что мы в состоянии воспроизводить наши мыслительные шаги сериально? Деннет отвечает на это так. Предварительно нужно задать другим вопросом: является ли различие в архитектурах мозга и компьютера теоретически важным различием? В одном отношении это различие неважно. Тьюринг доказал — и это, вероятно, самое большое его достижение, — что универсальная машина Тьюринга может вычислить любую функцию, которую в состоянии вычислить любой компьютер с любой архитектурой. Универсальная машина Тьюринга представляет собой некоего математического хамелеона, способного имитировать любые другие вычислительные машины: для этого ее только нужно снабжать подходящим описанием других машин. Подобно знаменитому Марселю Марсо — универсальной машине мимикрии человеческого поведения внешними приемами хореографии, — машина Тьюринга тотчас же начнет выдавать основанную на данном ей описании превосходную имитацию, и она превращается — виртуально — в другую машину. Поэтому компьютерную программу можно рассматривать либо как набор примитивных инструкций, которым надлежит следовать, либо как описание машины, которая должна быть имитирована.

Если сознание принять за виртуальную машину, или software, работающую на параллельном hardware мозга, что можно считать ее «программой» и как она действует? Гипотеза Деннета состоит в следующем. Имея глубокие биологические корни, виртуальная машина сознания тем не менее является *приобретенной программой* или системой программ. Правдоподобно предположение, что в какой-то момент очень далекого прошлого мозг человека «перепрограммировал» себя таким образом, что его генетически животное hardware стало «гнать» специфическую software, сформированную и нагруженную языком и культурой. То есть, поскольку синаптические связи мозга пластичны и изменение их конфигураций изменяет способы передачи информации, не исключено, что на каком-то этапе эволю-

156

ции человеческий мозг произвел их глобальную конфигурацию, в результате чего был создан существенно иной стиль движения информации, стиль, включающий в себя, по меньшей мере время от времени, последовательные, структурированные во времени, подчиняющиеся правилам виды деятельности, которые выглядят как типичные машины фон Неймана. То есть мозг стал работать сериально.

Ключевая идея этой гипотезы состоит в том, что, эксплуатируя пластику синаптических связей мозга, приобретенная программа, или система программ, одновременно проявляет свою собственную активность по отношению к нему: она *накладывает свой отпечаток на hardware*

*мозга* и ответственна за производство нового типа outputs. Здесь имеет значение следующее: у мозга имеется множество форм приспособительной пластики, которые могут выливаться в мириады различных микропривычек и оказывать влияние на различные макропривычки. В случае машины фон Неймана это выполняется сотнями тысяч нулей и единиц (битов), разделенных в зависимости от типа машины на «миры» в 8, 16, 32 или 64 битов, хранящихся в отдельных регистрах памяти и выдающих обозначенные в инструкции слова в определенные единицы времени. В случае параллельной машины мозга операции выполняются скорее всего тысячами, или миллионами, или миллиардами связующих устройств нейронов, которые в оркестровом исполнении снабжают hardware *новым набором макропривычек или новыми навыками поведения*. Имеется и другая дизаналогия между мозгом и hardware компьютера — различие в способе приобретения привычек. В случае машины фон Неймана программа с диска просто вводится в главную память компьютера, и он тотчас же обретает новый набор привычек. В случае мозга дело обстоит по-другому. Для того чтобы программы миллионов нейронных связей-усилений воплотились в органическом компьютере мозга, требуются *длительное обучение, тренинг*, повторяющаяся самостимуляция. Тренинг и обучение — важнейшие и необходимые условия. Пол Черчленд, считает Деннет, неправильно

157

его понял, приписав мысль, что виртуальная машина сознания является классической, состоящей из дискретных состояний и подчиняющейся правилам машиной. На самом деле «В машине фон Неймана вы только "вводите" программу с диска в главную память, и компьютер тотчас же получает новые привычки; с мозгом дело обстоит по-иному, ему требуется тренирующее обучение. ...Это, конечно, большая дизаналогия»<sup>1</sup>. Способ реакции на инструкции в фиксированном машинном языке Центрального процессуального блока (CPU) фон Неймана является жестким, и для сохранения его программ этот факт является определяющим. Человеческий мозг не является таковым. В нем нет ничего даже отдаленно напоминающего фиксированный машинный язык IBM или совместимых с IBM компьютеров. Вот почему мозги двух разных людей не имеют тождественной программы (это еще один аргумент, почему теория тождества духовного и телесного оказалась несостоятельной). В то время как одна и та же программа может быть запущена на компьютерах разных систем, нет оснований думать, что при сходном устройстве мозгов и при одинаковых условиях среды два сознания или две виртуальные машины будут иметь одну и ту же программу.

Итак, несмотря на дизаналогии, Деннет приводит несколько оснований, исходя из которых все же целесообразно апеллировать к компьютерной стратегии. Они таковы: 1) человеческое сознание — слишком недавняя инновация, чтобы стать жестким креплением во внутренней машине мозга; 2) сознание в большей мере является продуктом культурной эволюции, накладывающей свой отпечаток на мозг при обучении в раннем возрасте; 3) успешная его инсталляция в мозгу детерминирована мириадами микроструктур пластичности мозга, что объясняет, почему его функционально важные особенности невидимы при нейроанатомическом анализе и скорее всего никогда не будут видимы.

<sup>1</sup> Dennett D.C. and Densmore Sh. The Virtues of Virtual Machines. Dec. 20, 1996. Internet: <http://ase.tufts.edu/cogstud/pubpage.htm>. P. 1.

158

Самая важная дизаналогия, таким образом, состоит в том, что человеческое сознание — особая, «рожденная культурой виртуальная машина»<sup>1</sup>, и это накладывает отпечаток на все его проявления. Поскольку между двумя мозгами не может быть одного и того же машинного языка, способы передачи информации между виртуальными машинами сознания человека должны быть социальными, чувствительными к контексту культуры, способными к самоорганизации и самокорректировке. «Разговор» двух разных компьютеров, например систем Макинтош и IBM PC, обеспечивают инженеры. Разговор человеческих существ обеспечивают коммуникативные средства культуры, обучение, имитация поведения других людей, «подкрепления» и др.

Определяющая роль в развитии и совершенствовании виртуальной машины человеческого



сознания, конечно, принадлежит языку. Благодаря ему работающее на параллельном hardware мозга сознание приобретает сериальность и становится возможной его компьютерная имитация. Под «языком» понимается и разговорная речь, и письменность, и самые разнообразные знаковые и маркировочные системы. Подобно тому как использование колеса зависит от наличия мощных дорог или других искусственно спланированных поверхностей, виртуальная машина сознания работоспособна при определенных условиях: при наличии семиотических схем, коммуникации и обучения. Эти условия являются необходимыми, поскольку ее потребности в памяти и распознавании паттернов требуют, чтобы мозг «сбрасывал» некоторое количество его памяти в окружающую среду, т.е. в знаковые системы.

Одной из позитивных сторон аналогии сознания с виртуальными машинами, которыми оперируют компьютерные науки, по Деннету, является ее эвристичность. Когда на заре компьютерного века обсуждался дизайн компьютера, предполагалось, что это будет просто числовая машина. В настоящее

<sup>1</sup> Dennett D.C. Living on the Edge. Symposium: Daniel Dennett's «Consciousness Explained» // Inquiry. V. 36. № 1/2. 1993. P. 146.

159

время его числовые операции эффективно используются в создании новых виртуальных машин, мощь которых кажется почти магической. «Изначально мозг человека тоже не был спроектирован природой... для словесной процессуальности, сегодня же огромная порция — может быть, даже львиная доля активности, происходящей в мозгу взрослого человека, вовлечена в процесс словесного производства: в производство речи и понимания и в сериальную реорганизацию лингвистических единиц, или лучше — нейронных суррогатов. И эти активности магнифицируют и трансформируют лежащие в их основе силы hardware такими способами, которые кажутся («извне») совершенно магическими»<sup>1</sup>.

Поскольку, по Деннету, главная задача исследователя состоит в том, чтобы понять, как человеческое сознание реализуется в операциях *виртуальной машины*, созданной культурой в мозгу, соответствующим должен быть теоретический язык описания мультитрековой функциональной организации сознания. На этот счет у Деннета есть вполне определенное мнение: ни язык интроспекции, ни язык нейрофизиологии для этого не годятся; необходимый нам уровень теоретического объяснения и описания сознания аналогичен (хотя и не идентичен) тому, который называют «software level» в описании компьютеров.

Для корректного понимания позиции Деннета в отношении «виртуальности» важно иметь в виду два момента.

Во-первых, он оговаривает, что эксплуатация им специфического жаргона компьютерных наук («программа», «software», «hardware», «virtual machine» «drive» и др.) объясняется тем, что здесь появился словарь, который для описания сознания оказался более подходящим, нежели традиционный, вводящий в заблуждение, менталистский словарь. При этом следует помнить, что в философии сознания смысловая нагрузка терминов этого словаря зависит от специфического для нее контекста рассуждений и от прагматической применимости. Ска-

<sup>1</sup> Dennett D.C. Consciousness Explained. P. 225.

160

занное относится к эксплуатации биологических терминов и аналогий<sup>1</sup>.

Во-вторых, «виртуальность», о которой ведется речь, не следует понимать в смысле особого вида *реальности* (бытия или существования). Такую интенцию Деннет пресекает по той причине, что логически это ведет к воспроизведению — в новом облики — картезианского дуализма *res cogitans* и *res extensa*. Виртуальность не есть отдельное *свойство* или особый атрибут *наряду* с физическим миром, а только *функциональные отношения информационных потоков*. Придерживаясь жесткой антидуалистической стратегии, он с иронией относится к расплывшимся в самых различных дисциплинах и нещадно эксплуатируемым «виртуальным реальностям». «Развитие систем "Виртуальной Реальности", — пишет он, — сегодня переживает бум. Состояние этого мастерства впечатляет: здесь и электронные жесткие рукава, дающие

убедительный интерфейс (interface) для "манипулирования" виртуальными объектами, и ' В свете этой оговорки Деннета имеет смысл сделать ряд семантических пояснений. Начнем с применения к сознанию термина «виртуальная машина» (virtual machine). Со словом «машина» все более или менее просто: Деннет употребляет его не в смысле Ламетри («человек-машина»), а как *абстрактный* принцип организации. Сложнее обстоит дело со словом virtual. Мы переводим его калькой «виртуальный», не совсем удачной из-за многообразных значений этого слова в русском языке (возможный, потенциальный и др.). У Деннета прилагательное virtual по отношению к сознанию, как и термин «software program», тоже многозначен. С одной стороны, имеется в виду смысл, принятый в компьютерных науках для обозначения информационных процессов (не случайно, что свою теорию он чаще всего называет «AI процессуально-информационной моделью сознания»). С другой стороны, его «virtual» включает смыслы, принятые в повседневном английском языке — «фактический», «не номинальный», «действительный», «эффективный». Он хочет сказать, что хотя сознание и виртуально, но в то же время оно действительно как эффективно работающий механизм.

161

шлемообразные визуальные дисплеи (head-mounted visual displaces), предоставляющие возможность исследовать виртуальные условия окружающей среды большой сложности. Тем не менее ограниченности этих систем очевидны и подтверждают мою точку зрения: упорные иллюзии могут поддерживаться только благодаря различным комбинациям физических репликаций и схематизаций (относительно упрощенной репрезентации). И даже в лучшем случае они являются опытами виртуальной сюрреальности, а не чем-то, что вы могли бы на какой-то момент принять за реальную вещь. Если вам вздумается действительно обмануть кого-либо и заставить поверить, что он находится в клетке с гориллой, попробуйте прибегнуть к услугам актера в шкуре гориллы — в течение длительного времени это будет вашим лучшим аргументом в споре<sup>1</sup>. Другими словами, «виртуальные миры», которыми оперируют ученые, являются только физическими эффектами и как таковые находятся в пространстве и времени, в то время как виртуальная машина сознания, в которой мысли сами себя мыслят, находится «нигде». Эти разъяснения мы сделали для того, чтобы компьютерный жаргон не затемнял нам экспликацию метафизической линии рассуждений Деннета, а именно, как сознание мыслит, обходясь без центральной самости.

### 3. Эволюционизм, язык и роль меме-вирусов в производстве сознания

Как ранее говорилось, стратегию на применение к объяснению сознания «Машины Тьюринга» в совокупности с эволюционным подходом Деннет принял уже в своей первой книге «Содержание и сознание». Тогда она была еще достаточно абстрактной и со многими неизвестными. В дальней-

<sup>1</sup> Dennett D.C. *Consciousness Explained*. P. 6—7.

162

шем он попытался более детально разобраться в значимости эволюционной теории. Результатом явилась книга «Опасная идея Дарвина» (1995). Рост интереса к биологии сказался на применяемом им инструментарии: можно сказать, что он двигался от когнитивизма информационного толка к *эволюционному когнитивизму*.

Сегодня, говорит он, практически все называют себя натуралистами и эволюционистами. Эволюционные аргументы используются в эпистемологии, этике, политике и где угодно. Стала расхожей мысль о распространении эволюционных принципов (естественный отбор, выживание наиболее приспособленных, борьба за самосохранение и др.) на сферу научного знания, культуры, морали и т.п. Как это часто происходит со всякой здравой идеей, ставшей банальностью, эволюционный принцип оброс путаными интерпретациями: авторы без должной осторожности обращаются с дарвиновскими идеями. Часто забывается главное: эволюционизм — это только *общая абстрактная характеристика развития* — *abstractum*. Движущий фактор эволюции — естественный отбор — представляет собой понимание перспектив эволюции с помощью абстрактных паттернов. Это значит принятие установки дизайнера по отношению к эволюции,

акцентирующей внимание на организационные принципы функционирования и интенциональную установку в отношении бессознательной селекции. (Вспомним, что К. Поппер тоже говорил о дарвинизме как «метафизической исследовательской программе», а не теории, опирающейся на жесткие законы.)

Согласно Деннету, мы можем говорить, что эволюционный принцип действует, когда имеются следующие условия:

- 1) вариабильность: имеет место непрерывность изобилия (множественности) различных элементов;
- 2) наследуемость или воспроизведение себе подобных: у элементов есть способность и возможность создавать копии или реплики самих себя;
- 3) дифференциальная «приспособляемость»: число копий элемента, созданных в вариациях данного времени, зависит

163

от взаимодействия между признаками этого элемента (чем бы он ни отличался от других элементов), и особенностями окружающей среды, в которой он сохраняется<sup>1</sup>.

В этой общей и абстрактной характеристике эволюции ничего не говорится об органических молекулах, о питании, размножении и т.п., то есть специфически биологическом. Именно в силу этого эволюционизм как абстрактный принцип может использоваться универсально, и не только в биологии, но и социологии, психологии, в объяснении языка, культуры и компьютера. Последнее особенно примечательно.

В 60—80-х годах рядом ученых, занимающихся осмыслением AI, была высказана идея о том, что появление компьютеров — очевидных артефактов — следует рассматривать не только в рамках специфических закономерностей развития технической цивилизации, но и в рамках более общих биосоциальных процессов. Например, Марвин Минский, автор нашумевшей книги «Общество сознаний»<sup>2</sup>, проводит такую мысль: появление компьютеров следует расценивать как *следующий этап в общем эволюционном развитии*. Эта мысль близка мировоззрению Деннета, и к ней он обращался неоднократно: компьютеры — артефакты — не могут быть чужеродным образованием в теле Природы и не могут не вписываться в работу «Машины Дарвина». Самая трудная задача со многими неизвестными состоит в том, чтобы показать, или хотя бы высказать гипотезу, о возможных механизмах, благодаря которым «Машина Тьюринга» стыкуется с «Машиной Дарвина».

Для Деннета очевидно, что решение этой задачи нужно начинать с определения места и роли языка в эволюционном развитии. Ему импонирует идея, что естественный язык — самый важный и оригинальный артефакт из всех созданных

<sup>1</sup> Dennett D.C. Memes and the Exploitation of Imagination // The David and Miriam Mandel Lecture / American Society for Aesthetics. Oct. 27, 1989. Internet: <http://ase.tufts.edu/cogstud/pubpage.htm>. P. 1.

<sup>2</sup> Minsky M. The Society of Minds. N.Y., 1985.

164

человеком — в совокупности с нашим мозгом, являющимся его органическим носителем, можно расценивать как *первое изобретение универсального компьютера*. Эта идея в последнее время получила широкий резонанс в литературе. Философ Дж. Лейбер в книге «Приглашение в когнитивную науку» (1991) говорит следующее: «Рассматривая себя с компьютерной точки зрения, мы не можем избежать мысли, что естественный язык является нашим наиболее важным "программным языком". Это значит, что огромная часть знания и поведения выступает для нас в виде естественного языка и с его помощью лучше всего передается и понимается. Можно даже сказать, что естественный язык стал нашим первым великим и оригинальным артефактом. А поскольку языки являются машинами — этот факт осознается все больше и больше, — то естественный язык в совокупности с нашим мозгом, являющимся его носителем, стал нашим первым изобретением универсального компьютера. Все это можно было бы утверждать, хотя есть тайное подозрение, что язык не есть что-то нами изобретенное, а что-то, чем мы стали сами; он не

есть нечто, нами сконструированное, а что-то, в чем мы творим себя и заново сотворяем себя»<sup>1</sup>. Это «подозрение» стыкуется с витгенштейнианской посылкой Деннета о сотворении человека языком. Но в отличие от Витгенштейна он натурализирует этот посыл.

Ряд важных подсказок для детализации стратегии объяснения сознания совместной работой «Машины Тьюринга» и «Машины Дарвина» Деннет нашел в работах социобиологов Эдварда Уилсона («Sociobiology: The New Synthesis», 1975), Ричарда Докинса («Эгоистичный ген», 1974; «The Extended Fenotype», 1982) и в личных беседах с этими учеными. В книге «Опасная идея Дарвина» (1995) он вторгся в нефилософские области и высказал свое видение эволюционной теории. Вольно и невольно он оказался вовлеченным в «Дарвиновские войны», ведущиеся в англоязычной мысли, в которых, поми-

<sup>1</sup> Leiber J. Invitation to Cognitive Science. Oxford. 1991. P. 8.

165

мо биологов, участвуют психологи, эволюционные эпистемологи и представители других специальностей. Предметом острого идейного конфликта являются ключевые для дарвинизма темы: понимание механизмов эволюционного развития, роль естественного отбора, адаптации, случайностей, генетических изменений и др. В этом конфликте он находится на стороне «адапционистов». Теория генно-культурной эволюции Ричарда Докинса, опирающаяся на принцип естественного отбора, как нельзя лучше соответствует изначальным его установкам. Оппозиционная сторона представлена многими именами, среди которых лидируют Стивен Гоулд и Ричард Левонтин. Последние утверждают, что многие продукты эволюции являются побочными эффектами, следствиями ограниченностей развития либо случайностей. Например, причина исчезновения динозавров и последующего распространения млекопитающих скорее всего кроется не в действии естественного отбора, а в случайности — падении астероида. В свете деннетовского понимания эволюции как абстрактного принципа эмпирический пример случайности, конечно, не есть весомый аргумент.

Деннетом была принята на вооружение важная идея социобиологов: с появлением человека возник новый вид репликаторов — культурные единицы. Докинс придумал для них имя — *меме* (слово memes имеет греческие корни, происходит от «mimeme»). Меме — это не простые «идеи» Локка и Юма, вроде ощущений и восприятий, а сложные образования, оформляющиеся в запоминающиеся семантические единицы, такие как идеи колеса, носильной одежды, алфавита, календаря, исчислений, прямого угла, шахмат, импрессионизма, деконструктивизма и т.п. Они выступают единицами культурной трансмиссии или единицами имитации, образуя веревки информационных кодов. Интуитивно они воспринимаются в виде более или менее определенных мысленных комплексов, фигурирующих в самых разнообразных областях культуры. Главное, что они способны к саморазвитию и эволюции. Об эволюции меме, оговаривает Деннет, следует говорить буквально-

166

но, а не фигурально или метафорически: как и гены, они подвержены действию *эволюционных процессов, конкуренции и естественного отбора*. Многие из них — однодневки, исчезающие, не оставив следа, а многие выживают; они выживают в том случае, если становятся единицами репликации или единицами имитации при адекватном использовании окружающей среды. В конкурентной борьбе они занимаются самоутверждением и самопропагандой и переходят, как семя от семени в биологическом виде, от одного человеческого мозга к другому. В процессе репликации отнюдь не всегда участвуют благие меме. Как и в случае с компьютерными вирусами, могут плодиться меме, лишенные всякого функционального оправдания, могут возникать эпидемии зловредных вирусов-меме, с готовностью паразитирующих на человеческих мозгах, лишая жертвы инфекции соревновательных преимуществ. Тому пример — история веками воспроизводимых суеверий, мифологических идеологий и т.п. Ричард Докинс был прав, утверждая, что некоторые признаки сознания могут быть именно «эгоистичными генами»<sup>1</sup>.

Нужно исходить из того, говорит Деннет, что в дизайн человеческого сознания внесли свой вклад три медиума — генетическая эволюция, пластичность фенотипа и эволюция мемосферы. В

сравнении с генотипом и фенотипом эволюция меме является сравнительно недавно возникшим медиумом и ограничена только видом *homo sapiens*, тем не менее роль эволюции меме для возникновения сознания оказалась решающей. Более того, появление популяции меме или, как он ее называет, «Мемосферы», следует рассматривать как новый этап эволюционного развития после революционных изменений, вызванных появлением генотипа и фенотипа. Как и всякие эволюционные новации, мемосфера оказывает обратное воздействие на породившие ее силы. Популяция меме явилась эволюционной новацией не только в смысле появления возможностей для творчества артефак-

<sup>1</sup> Dennett D.C. *Consciousness Explained*. P. 221.

167

тов, но и в смысле внесения изменений в дизайн самой природы. Одна из наиболее смелых гипотез Деннета состоит в следующем: меме влияют на архитектуру человеческого сознания и мозг, делая восприимчивым к ним это физическое тело. Характер используемой знаковой системы, способ обучения и имитации оказывают хотя и невидимое для нейроанатома, тем не менее заметное *воздействие на архитектуру мозга и его* компетенцию. Вирусы-меме «паразитируют» на нем и запускают виртуальную машину, превращая мозг в генератор сознания<sup>1</sup>.

Все это возможно потому, что подобно вирусам меме постоянно находятся в поиске подходящей среды для заселения. «Небеса, на которые рассчитывают попасть все меме, — это человеческое сознание, однако человеческое сознание само есть артефакт, сотворенный тогда, когда меме реструктурировали человеческий мозг для превращения его в лучшую обитель для меме»<sup>2</sup>. «Как только в мозгу выстроились входные и выходные пути для колесниц языка, они вскорости были *паразитированы* (я говорю об этом буквально) сущностями, которые сами эволюционировали, для того чтобы захватить эту нишу: меме»<sup>3</sup>. Трансформация человеческого мозга под влиянием вирусов-меме, говорит Деннет, является самым решительным эволюционным изменением в компетенции этого органа. «Тысячи меме, большинство из которых рождены языком, но также бессловесными "образами" и другими структурными данными, получают резиденцию в индивидуальном мозгу, формируя его тенденции и превращая его таким образом в сознание»<sup>4</sup>. Он высказал предположение, что специфические семиотические системы разных народов как-то влияют на архитектуру мозга и на ментальность. Гипотеза об активном воздействии мемосферы на органику мозга объясняет различия в ментальности

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 210.

<sup>2</sup> Ibid. P. 207.

<sup>3</sup> Ibid. P. 200.

<sup>4</sup> Ibid. P. 263.

168

народов, имеющих и не имеющих письменности. «Дописьменная ментальность, — говорит он, — вполне возможно включает значительно иной класс виртуальных архитектур по сравнению с теми, с которыми имеют дело грамотные общества»<sup>1</sup>. Или еще. Разница в мозгах людей, чей родной язык китайский, а не английский, объясняет различие в компетенции этих людей, что непосредственно наблюдается в поведении и проявляется во многих экспериментальных контекстах с носителями разных языков. Не исключено, что эти функциональные различия физически воплощены в микроскопических изменениях мозга. Они сегодня невидимы для нейроученых, а скорее всего никогда не будут увидены, какими бы тонкими ни были эксперименты.

В свете высказанных Деннетом гипотез понятно, почему он считает ментальный и нейрофизиологический языки непригодными для описания динамичной, восприимчивой к меме-вирусам архитектуры мозга. Для описания функциональной архитектуры сознания, *сотворенной* наводнением меме, прежний язык и методология, говорит он, не годятся; требуется нетрадиционный уровень описания. Такой уровень описания и объяснения аналогичен (но не идентичен) тому, который называют «software levels» в описании компьютеров. «Человеческое сознание *как таковое* представляет собой огромный комплекс меме (точнее, меме-эффектов в



мозгу), который лучше всего понимать как операции виртуальной машины типа «фон Неймана», *имплементированные в параллельной архитектуре* мозга, дизайн которой не был создан для такого рода активностей. Силы этой *виртуальной машины* в огромной степени увеличивают лежащие в основании силы органической *hardware*, на базе которой они действуют...»<sup>2</sup>. Представление сознания в виде машины фон Неймана не вступает в противоречие с эволюционным принципом. Сторонники сильного варианта AI, говорит он, «знают, что их базовые теорети-

<sup>1</sup> Ibid. P. 220.

<sup>2</sup> Ibid. P. 210.

169

ческие рамки есть не что иное, как прямолинейное распространение — на человеческий мозг и на весь его периферийный аппарат и интерфейсы — дарвиновской программы бессознательного механизма, проделывающего, в конечном счете, всю работу»<sup>1</sup>.

Подведем краткие итоги концепции Деннета о взаимодействиях «Машины Тьюринга», «Машины Дарвина» и «Машины Сознания». В данной главе мы постарались оттенить социобиологическую парадигму, в которой он работает, динамику сопряженности и взаимовлияния природных и культурных факторов. Как мы уже сказали, он рисует картину сознания в *дуальной перспективе* (в отличие от дуалистической). С одной стороны, он поднимает человеческое сознание над природной средой и животной психикой, утверждая, что стиль его информационной организации принципиально отличен от информационной организации животных, но сходен с компьютером. С другой стороны, он подчеркивает непрерывность сознания с животной психикой и глубинными эволюционными алгоритмами, в том числе диспозициями к обучению. «Обучение должно быть эволюцией в мозгу, следовательно, мозгу должны быть присущи алгоритмы обучения»<sup>2</sup>. Эта дуальная перспектива, как мы увидим далее, видна во всем — в его трактовке сознания, самости, свободы воли, морали и ответственности. Она проявляется и в применяемых им когнитивных средствах — *illata* и *abstracta*, интенциональное и реальное, виртуальное и действительное и др. Деннета часто упрекают за то, что, элиминировав из своей теоретической схемы феноменальное, он представил человеческого субъекта в виде зомби или чисто когнитивной машины. Нам же кажется, что деннетовский субъект похож не столько на зомби, сколько на киборга, т.е. на существо, сращенное из двух суб-

<sup>1</sup> Dennett D. In Darwin's Wake. Where I am. Dec. 29, 2000. Internet: <http://ase.tufts.edu/cogstad/pubpage.htm>. P. 3.

<sup>2</sup> Из интервью Д.Деннета. 07. 02. 2000. Интернет: <http://www.theatlantic.com/unbound/digitcult/dennett.htm>. P. 1.

170

станций — биологической и социальной, интеллект которого возник в результате их сложного взаимодействия.

Дуальная перспектива, фокусирующая динамику разных факторов, с нашей точки зрения, привлекательнее в сравнении с односторонним биологизмом и социологизмом. И все же она не снимает присущие деннетовской схеме напряженности. Они фиксируются его коллегами и оппонентами, тоже работающими над разгадыванием загадок сознания. Начнем с того, что без ответа остается вопрос о возможности сравнения компьютеров, оперирующих синтаксическими знаковыми системами, с человеческим сознанием, оперирующим семантическими единицами. Это постоянный предмет критики со стороны его оппонентов. (Более подробно об этом мы будем говорить в главе IX.)

Энди Кларк, близкий по мировоззрению и методологии к Деннету философ и известный специалист по робототехнике и искусственной жизни (Allife), считает гипотезы Деннета, безусловно, привлекательными и, что самое главное, стимулирующими мысль и новые поиски. Тем не менее он видит в них ряд болевых моментов, которые не смогла устранить ни одна из существующих теорий. Прежде всего, остается открытым вопрос, является ли роль языка, культуры и обучения решающей для возникновения сознания. Или эволюцией уже были созданы

предпосылки для этой роли — нейронные инновации, которые в дальнейшем были оформлены языком и культурой как человеческое сознание? То есть, делает ли нас существами самосознающими культурная «добавка» или для усвоения этой «добавки» должна существовать биологическая база? Остается открытым другой, связанный с первым вопрос — происхождение языка и вовлечение его в общий динамический процесс. Ведь ни одно животное в мире не в состоянии выразить информационные взаимодействия с телом и средой в форме, по глубине, широте и выразительности сравнимой с лингвистическим кодированием. Или еще одно недоумение: как это возможно, что плетение нарративов делает *бытийственным* человеческое сознание? Деннет предложил

171

несколько увлекательных идеи и гипотез, но ни одну из них, считает Кларк, нельзя принять как полностью удовлетворительную. Самый трудный вопрос: как именно инструменты сознания — язык и культура — осуществляют свою магию? С горизонта и его, и моей схемы, говорит Кларк, не сходит вопрос: «Как в деталях они работают?» «Когда он встает во весь рост, ни у Деннета, ни у меня нет ничего конкретного, что можно предложить, кроме нескольких импрессионистских спекуляций»<sup>1</sup>. До тех пор, пока не будут известны детали работы этой схемы, вряд ли можно быть уверенным в ее правильности. Сам Кларк не исключает предположения, что публичный язык не был прямой необходимостью для мышления. Наоборот, он стал случайной ветвью, на которой у человеческих существ возникало мышление. Во всяком случае, эксперименты с шимпанзе, которых учили решать проблемы (определять тождественность и различие у предметов с помощью бирок и наклеек), свидетельствуют о весьма сложном уровне их мышления и какой-то способности к абстракции. Все это делает правдоподобной гипотезу о наличии у них биологических базисных форм обучения и понимания.

Примерно в том же духе высказывает сомнения Пол Черчленд, который, по его словам, плывет на том же материалистическом корабле, что и Деннет, и тоже является «скрытым гегельянцем» в том отношении, что стремится нарисовать «динамический профиль» концептуальной эволюции, но избрал другой путь решения проблемы сознания. Грубую ошибку Деннета он видит в том, что тот заимствовал у Докинса идеи «меме» и генно-культурных аналогов. Отсюда неправдоподобность его гипотезы, что специфически человеческое сознание возникло в результате паразитирования меме на материале мозга, отчего тот приобрел новую программу и новый стиль — дискурсивный, сериальный, аналогичный сти-

<sup>1</sup> Clark A. That Special Something: Dennett on Making of Minds and Selves // Daniel Dennett. Ed. by Andrew Brook and Don Ross. Cambridge. 2002. P. 191.

172

лю работы машины фон Неймана. Поэтому он предлагает отбросить разговор о «меме-вирусах», а заодно и все компьютерные метафоры («software», «hardware», «virtual machine», «drive» и др.).

Мозги животных и людей, говорит Черчленд, не являются машинами фон Неймана. Их кодирование не является цифровым, происходящие в них процессы не являются сериальными, они не выполняют накопленные в них программы и у них нет регистров накопителей. Опыт пятидесяти лет нейронауки и пятнадцати лет нейромоделирования учит, что мозг представляет собой «котел» совершенно иного рода. Вот почему он столь беспомощен при умножении в уме двух двадцатизначных чисел — задачи, которую компьютер делает за секунду. И вот почему компьютеры столь беспомощны в решении других задач, таких как распознавание индивидуальных лиц или понимание речи, что мозг делает мгновенно. Деннет прекрасно все это знает, и это не может не составлять для него проблему.

Как и Кларк, Черчленд исходит из того, что феномен сознания представляет собой результат базисных hardware структур — структур, которые широко распространены во всем животном царстве. Ключ к пониманию специфической ткани когнитивности, обобщенно именуемой «сознанием», лежит в понимании динамических свойств биологических нейронных сетей с высокочастотной физической архитектурой — архитектурой, которая репрезентирует hardware, присущий мозгам всех животных, а не только уникальные «software» черты человеческих мозгов<sup>1</sup>.

<sup>1</sup> Churchland P.M. Catching Consciousness in a Recurrent Net. Op cit. P. 65.

## V. Самость как «центр нарративной гравитации» или почему возможны самостные компьютеры

В мозгу не существует клетки или группы клеток, обладающих таким анатомическим или функциональным преимуществом, чтобы играть ключевую роль или быть центром гравитации всей системы.

*Уильям Джемс*

Подлинная человеческая самость по большей части сотворена бессознательным интерперсональным процессом дизайна, в который мы вовлекаем маленьких детей, чтобы превратить их в коммуникаторов...

*Дэниел Деннет*

Если в моей голове обитает сообщество компьютеров, тогда здесь, по идее, должен быть кто-то, кто заведует ими; ей-богу, было бы лучше, чтобы это был Я.

*Джерри Фодор*

### 1. Биологическая самость и нарративная самость

Что такое самость? Каково ее онтологическое место в мире? Как соотносятся мое Я и мое тело? Со времен Декарта такими вопросами задавались философы, и в их ответах самости всегда придавался уникальный статус. Считалось, что самость (или «Я», или «душа», или «эго») составляет суть индивиду-

174

ального бытия человека, особое качество, отличающее его от других людей, от животных, от физических вещей. В сложившейся после Декарта традиции самость обычно помещается в *центр* активности сознания: предполагается, что она стягивает к себе информацию, идущую из органов чувств, анализирует ее, обобщает и дает команды действиям человека. Ей приписывается качество самотождественности, обеспечивающее *метафизически личностное* единство индивида от рождения до смерти. Важнейшим свойством самости считается *самоактивность*; в ней усматривается смыслообразующий источник, из которого произрастает интимно-личностное освоение мира и все многообразие культуры. Атрибутом самоактивной самости считается *свобода воли*, в свою очередь выступающая основанием целеполагания и *моральной ответственности*. Правда, были и скептики, сомневавшиеся в существовании самости. Юм, например, говорил о невозможности выделить то, что кажется собственной самостью, о ее удивительной ускользаемости от анализа.

Дэниел Деннет продолжил и существенно радикализировал юмовский скептицизм. Представление о запрятанной в каждом из нас некоей драгоценной жемчужине — самости — возникло, говорит он, из-за приверженности дефективному набору картезианских метафор и неоправданного доверия к собственной интуиции. Такой реальности нет в природе: самость — мнимый режиссер Картезианского Театра. На самом деле — это воображаемый центр, который мы сами сплетаем из собственных историй и рассказов о нас других людей. Поэтому объяснение самости (и сознания) следует перевести с традиционного языка интуиции на совершенно иной уровень — языковый и информационно-процессуальный — и рассматривать ее в виде деятельности виртуальной машины, локомотивом которой является природно-культурная эволюция. Вполне правомерно предположение, что соответствующе запрограммированный робот в принципе мог быть сознательным и считаться самостью.

175

Такой подход к самости представлен в его книгах «Пространство для свободы воли», «Интенциональная установка», «Сознание объясненное» (гл. 13), «Свобода развивающаяся», в статьях «Почему каждый из нас является новеллистом» (перевод которой представлен в Приложении 1), «Происхождение самостей» и в других работах<sup>1</sup>.

Почему же человеку все-таки свойственно стойкое представление, что его самость реальна, что

она существует где-то внутри его? На этот вопрос у Деннета имеется эволюционный ответ, вернее, два типа эволюционного ответа. Первый ответ натуралистический. У самоощущения нашего Я имеется мощная биологическая основа. Возникновение у животных адекватного поведения в среде было связано с появлением способности к различению между «мною» и «остальным миром». Даже амеба бессознательно должна проводить такое различение для своего самосохранения. Минимальную склонность отличать себя от всего другого можно назвать *биологической самостью*. Под биологической самостью, оговаривает Деннет, не следует понимать что-либо конкретное или особую реальность, это только *абстракция, функциональный принцип организации эволюции*. Данный принцип организации присутствует в поведении птиц, бобров, термитов, пауков, бессознательно включающих жизненную территорию в свое «Я». Причем для их «самостности» вовсе не требуется «сознательность», «знание» или «мудрость».

Самые необычные принципы функциональной организации созданы приматами — homo sapiens. Каждый нормальный индивид этого вида занят творчеством самости. Из своего мозга он вытягивает паутину слов, и, как паутина пауку, они обеспечивают человеческому индивиду жизнеспособность и коммуникацию в среде. «"Паутина дискурса"... является та-

<sup>1</sup> Elbow Room: the Varieties of Free Will Worth Wanting. Oxford, 1984; The Intentional Stance. Cambridge, L., 1989; Dennett D. Why Everyone is a Novelist // Times Literary Supplement. Sept. 16—22. 1988. P. 96—98; The Origins of Selves // Cogito, 3, 1989. P. 163—173. Freedom Evolves. N.Y., 2003.

176

ким же биологическим продуктом, как и любые другие конструкции, которые мы обнаруживаем в животном мире»<sup>1</sup>. Человеческая тактика самосохранения, самоконтроля и самоутверждения, говорит Деннет, состоит в нарративной деятельности, в придумывании сюжетов и осуществлении контроля над историями, которые мы рассказываем и себе и другим людям о том, что мы есть. При этом важно следующее: подобно тому как бобрам для строительства плотин не нужно обладать знанием о плотинах или проводить их проект через сознание — они просто заняты их строительством, — человеку в плетении языковых нарративов не нужно *знать*, что он при этом делает, или проводить через «сознание»: он просто делает это. Потоки нарративов исторгаются, *как если бы* из единого источника — не только в физическом смысле проистекания из одного рта или одного карандаша, но в более тонком смысле; они воздействуют на аудиторию, побуждая слушателей постулировать единого агента, чьими словами они являются, и о ком они говорят. Некоторые из производимых речевых набросков могут содержать такие, например, высказывания: «в центре моего мира находится мое Я», «мое Я связывает мое детство и взрослую жизнь». Такого рода речевые построения порождают в нас интуитивное чувство, что где-то в центре нас находится «самосознающая самость», что именно она является источником значений и смыслов, окрашивая опыт с помощью квалиа. Одним словом, интуицию, подсказывающую, что самость является скрытой внутренней структурой помимо того, что раскрывают нам исследования функционирования мозга, речевых продуктов и внешнего поведения. Все эти интуиции, говорит Деннет, следует относить к иллюзиям Картезианского Театра. Утверждать, что только человеческий мозг и сознание творят самость, — ошибочно. Некоторые философы и ученые всерьез озабочены идеей найти в эволюции точку, где биологическая самость «вдруг» превратилась в человеческую самость, способную выйти за

<sup>1</sup> Dennett D. Consciousness Explained. L., 1991. P. 416.

177

пределы запрограммированных эволюцией образцов поведения и начать плести веревки слов и знаковую культуру. Искать такую точку, говорит Деннет, бессмысленно, это все равно как искать «точку» возникновения млекопитающих. Утверждать, что, сотворив мозг, эволюция сотворила и самость, ошибочно. В какой точке эволюционной истории появились разумно-понимающие, реальные самости? Не задавайте такой вопрос, будет большой глупостью пытаться идентифицировать первый или наиболее простой случай такой «реальности»<sup>1</sup>.

Самость есть артефакт, возникший тогда, когда язык и культура реструктурировали человеческий

мозг, сделав его способным оперировать культурными единицами — меме. «Наши рассказы о себе плетутся, однако по большей части не мы их сплетаем; они плетут нас. Наше человеческое сознание и наша нарративная самость представляют собой результат, а не их источник».

Для обозначения притягательности индивидуального сознания к дискурсу и меме Деннет придумал (перефразируя У. Джемса) метафору «*центр нарративной гравитации*», хорошо передающую присутствие и в то же время виртуальность самости. В статье «Почему каждый из нас является новеллистом» он предлагает читателю поучаствовать в мысленном эксперименте, а для этого рассмотреть несколько уместных для самости аналогий<sup>2</sup>.

Первая аналогия — с физическим центром гравитации. Физический центр гравитации является столь же загадочным явлением, как и самость, и обладает некоторыми общими с ней свойствами. В физике Ньютона центр гравитации объекта не есть какая-то частица вещества или реальный, дополнительный к атомам элемент Вселенной. У него нет никаких физических свойств, кроме пространственно-временной локализации и прерывного перемещения. Это чисто абстрактный объект или изоб-

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 418.

<sup>2</sup> Dennett D. *Why Everyone is a Novelist // Times Literary Supplement*. Sept. 16—22. 1988. Деннет Д. Почему каждый из нас является новеллистом. См. перевод в Приложении 1 данной книги. С. 428—447.

178

ретенная теоретиком фикция (fiction). В этом отношении он может служить примером того, что Ганс Рейхенбах назвал бы *abstractum*. Тем не менее в рамках физики роль этой фикции тщательно изучена, определена и хорошо работает. Если же пытаться установить каузальную зависимость между центром гравитации и каким-нибудь физическим элементом, будет допущена серьезная категориальная ошибка, ибо центр тяжести есть *не что иное*, как *abstractum*. Он представляет собой не *illata*, то есть выводимую сущность, а только воображаемый, не привязанный к чему-то конкретному объект.

Самость тоже следует рассматривать как абстрактный объект и порождение теоретика, не привязанное к какому-либо конкретному носителю. Физики знают, что они прибегают к очень большому упрощению, постулируя для какого-либо объекта центр гравитации или единую точку, по отношению к которой исчисляются все гравитационные силы. Гетерофеноменологи также отдают себе отчет в том, что производят чрезвычайное упрощение, используя абстракцию «центр нарративной гравитации» применительно к плетущему нарративы человеческому телу. Психологическая или нарративная самость, как и биологическая самость, не есть какая-то вещь в теле или мозгу, а только другая абстракция или принцип функциональной организации. Тем не менее эта абстракция является почти осязаемым аттрактантом свойств и «владельцем записей» о любых окружающих тело вещах и явлениях.

Интересные виртуальные аспекты самости высвечиваются при сравнении ее с другой разновидностью воображаемых объектов — вымышленными персонажами литературы. Многие придуманные писателями герои получают новую жизнь после публикации: можно рассуждать об их судьбе, экстраполировать их ситуацию, приписывать возможные мысли и т.д., то есть заниматься литературной интерпретацией текстов. Интерпретация и экстраполяция создают огромные возможности для манипуляции с воображаемыми объектами. В пост-публикационной жизни воображаемых героев есть одно важное сходство с другими воображаемы-

179

ми объектами: у них отсутствует жесткая каузальная связь с носителями. Имеет место во многом автономное (содержательное) движение информационного потока. Подобно тому как физические центры гравитации не привязаны к какой-либо атомной частице вещества, вымышленные литературные персонажи не привязаны к телу или автобиографии автора, или к обстоятельствам, которые послужили ему основой для их придумывания. Самость, будучи нашей собственной воображаемой автобиографией, тоже не привязана к какой-либо «реальной», «подлинной» самости или к какому-нибудь нейрону мозга (искать самость в точке мозга Деннет считает очень



серьезной категориальной ошибкой). Критикам, утверждающим, что его отказ от метафизически реальной самости в пользу нарративной самости на самом деле есть ее элиминация, Деннет отвечает, что речь идет о переводе разговора о самости на другой уровень: мы есть то, о чем мы говорим, мы есть то, что говорят о нас. «Вы есть то, над чем вы осуществляете контроль. Это так просто»<sup>1</sup>.

Отказ от образа самости-режиссера, командующего сценическими действиями Картезианского Театра, и принятие метафоры центра нарративной гравитации, естественно, потребовало от Деннета пересмотра всех привычных самостных характеристик. Например, *сколько у человека самостей?* Люди привыкли думать, что каждому человеку, точнее, каждому человеческому телу, соответствует одна самость. Научные исследования последних десятилетий ряда аномальных явлений, говорит он, показали ошибочность такого представления. Например, выводы, сделанные на основе исследований синдрома множественной личности (MPD), пациентов с рассеченными полушариями мозга, людей, у которых отсутствуют ощущения, а также ряда других отклонений от нормы, заставляют серьезно усомниться в этом общепринятом мне-

<sup>1</sup> Dennett D. Living on the Edge. Symposium on Daniel Dennett's «Consciousness Explained» // Inquiry. V. 36. № 1/2. 1993. P. 145.

180

нии<sup>1</sup>. В частности, в полученных результатах о феномене MPD он видит подтверждение его гипотезы, что единой самости, привязанной к одному телу (мозгу), не существует. У некоторых людей, перенесших психологическую травму, иногда происходит раскол личности. Сохраняя себя, они как бы «утекают» в другие воображаемые личности и творят другие самости. Все это скорее свидетельствует не о существовании некоей первичной амебоподобной самости, приспособляющейся к обстоятельствам, а о том, что разные нарративы формируют разные самости. Есть случаи, когда два разных тела имеют одну самость (близнецы Грета и Фреда в Великобритании с разными телами говорят в унисон, одна заканчивают фразу, начатую другой, у них одни и те же предпочтения, вкусы и др.). Много споров возникло в связи с операциями по рассечению мозга (при заболеваниях эпилепсией и др.). Некоторые философы полагают, что разъединение полушарий мозга создало две самости. На самом деле, говорит Деннет, происходит другое: физическое вторжение в мозг изменило вербальные условия рассказывания субъектом своих историй. Одним словом, глубоко засевшее в нас представление «одно тело = одна душа» отнюдь не является аксиомой.

Столь же ложным, с его точки зрения, является бытующее представление о *самотождественности и длительности самости*. В последнее время философы много дров наломали по этому поводу. Являюсь ли я той же самой личностью, чьи приключения в детском саду иногда живо всплывают в моих воспоминаниях? Считается, будто память, эта наиболее характерная черта ментальности, предполагает, что у помнящей личности на протяжении жизни сохраняется активное чувство постоянства самости. Представление о себе как о неизменных субъектах опыта, иммунных к телесным и нарративным изменениям и к изменениям социальной среды, является иллю-

<sup>1</sup> Деннет вместе с Н. Хэмфри специально занимался исследованием людей с MPD: Humphrey N. and Dennett D. Speaking for our Selves: An Assessment of Multiple Personality Disorder // Raritan, 1989. № 9.

181

зией<sup>1</sup>. Чувство самотождественности и личностного постоянства создается постоянством обстоятельств. Философ Дирек Парфит (1984) сравнил личность с клубом, который однажды может прекратить свою деятельность, а затем возобновить ее. Является ли он тем же самым клубом? Да, говорит Деннет, в том случае, когда у него тот же самый устав, то же название и функции, те же члены и т.д. Сказанное относится и к самостям. «Самости — это не независимо существующие души-жемчужины, а артефакты творящих нас социальных процессов. Подобно другим таким же артефактам, мы являемся субъектами с резко меняющимися статусами. Единственная «инерция», точно подогнанная к траектории самости, или клуба, — это

стабильность, которая приписывается ей паутиной конституирующих ее верований, и когда эти верования рушатся, самость тоже рушится либо перманентно, либо временно»<sup>2</sup>.

## 2. Почему возможно помыслить самостный компьютер

Итак, самость, по Деннету, — это человеческое бытие, сплетенное из нарративных веревок. Мы рисуем самих себя и других людей с помощью наших собственных и произведенных другими людьми дискурсов, собирая их в более или менее

<sup>1</sup> Нужно сказать, что особый социо-лингвистический (деятельностный) подход к самоидентификации личности развивали советские психологи. А.Н. Леонтьев в книге «Деятельность. Сознание. Личность» (М., 1975) писал: «Вопреки своей распространенности, взгляд на личность как на продукт биографии человека является неудовлетворительным... Упускается главный психологический факт, а именно, что человек вступает в отношение к своему прошлому, которое по-разному входит в наличное для него — в память личности» (с. 217).

<sup>2</sup> Dennett D. *Consciousness Explained*. P. 423.

182

когерентный узел. Привычная идея Я, как реального и единого агента или центра, вполне уместна и полезна для психологических, социальных и классификационных целей, то есть имеет только прагматический, но не теоретический статус.

Обратимся снова к статье Деннета «Почему каждый из нас является новеллистом» и посмотрим на продолжение его мысленного эксперимента. Наряду с аналогией между самостью и физическими центрами гравитации, самостью и воображаемыми персонажами литературы, он предлагает порассуждать о правдоподобности аналогии между самостью и компьютером. Любопытный поворот в мысленной игре с самостью произойдет, если вообразить пишущий повести компьютер. Компьютер является простой, стоящей в лаборатории машиной. Начиненный им робот, которому дизайнеры дали имя «Гилберт», не что иное, как зомби, который не ведает, что творит (наш мозг, напоминает Деннет, тоже не ведает, что творит). Предположим, что снабженный всевозможной информацией робот способен еще путешествовать по миру и реагировать на него. Имея внутри себя компьютер с соответствующей программой, он в принципе мог бы выдать на принтер повесть с таким, например, названием: «Жизнь и времена Гилберта». Несведущему посетителю библиотеки будет трудно отличить ее от повести, написанной живым человеком. Критический для эксперимента вопрос состоит в следующем: «Кто такой Гилберт?» Является ли Гилберт просто роботом или созданной роботом воображаемой самостью? Деннет склоняется к следующему варианту ответа. «Контролируемый компьютером поведенческий паттерн в принципе может быть интерпретирован нами как разрастающаяся биография или нарративное повествование о самости. При этом мы — не единственные интерпретаторы. Робот-рассказчик тоже, конечно, является новеллистом и интерпретатором; самоинтерпретатором, выдающим собственное объяснение своей деятельности в мире»<sup>1</sup>.

Можно рассуждать о достоверности или недостоверности этого повествования, однако робот «Гилберт» тоже может быть

<sup>1</sup> Деннет Д. *Почему каждый из нас является новеллистом*. С. 436.

183

интерпретатором и даже самоинтерпретатором, выдающим собственное объяснение своему путешествию в мире. Поэтому вполне логичен такой вывод:

«Если самость есть "именно" "центр нарративной гравитации", и если все феномены человеческого сознания объяснимы "именно" как активности виртуальной машины, реализуемой в астрономически приспособительных связях мозга, тогда в принципе соответствующе "запрограммированный" робот с компьютерным мозгом на силиконовой основе был бы сознательным, был бы самостью. Говоря точнее, могла бы существовать сознательная самость, чьим телом был бы робот, а мозгом компьютер»<sup>1</sup>.

Разумеется, речь идет о гипотезе. Пока реальный и достаточно софистичный робот «Ког», над которым работают специалисты из Массачусетского технологического института (в проектах

которого принимал участие Деннет), еще не «Гилберт», и он далек от признаков сознательности и самостности. Тем не менее, признает Деннет, когда глаза «Кога» хотя и тупо, но следят за движениями человеческой руки, а его кисти гибко повторяют ее движения, трудно подавить в себе чувство, что рядом находится другой сознательный наблюдатель.

Итак, допустим, что в принципе возможен компьютер с «центром нарративной гравитации». Но как быть с «квалиа», связанными с интимно-эмоциональной стороной жизни, которые обычно соотносят с самостью? Можно представить умело манипулирующего конечностями, софистичного и даже пишущего автобиографические повести робота, но представить эмоционального робота трудно. Играющие в шахматы компьютеры могут превзойти по хитроумности живых игроков, но вряд ли они будут в порыве эмоций останавливать часы, сбрасывать фигуры с доски и оскорблять друг друга. Сотворенный компьютерным сердцем робота персонаж «Гилберт» может сказать «мне больно», если по конструкции робота ударят палкой, но все-таки это не та же самая реакция, кото-

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 431.

184

рая бывает у человека и животного, когда их бьют палкой. Если согласиться с Деннетом, что расцвеченный всевозможными квалиа — вкусами, болью, запахами — «феноменальный сад» не что иное, как мираж, то как объяснить факт, что именно человеку, а не роботу, свойственно интимно-личностное и эмоциональное переживание своего бытия, что он восхищается симфонией Бетховена и негодует по поводу нанесенного ему оскорбления?

Деннет прекрасно понимает, какое внутреннее сопротивление у человека вызывает «дисквалификация квалиа». Он вспоминает фразу сомнения, сказанную ему в беседе за бутылкой вина Уилфредом Селларсом, философом, много сделавшим для критики феноменализма: «Но, Дэн, ведь квалиа как раз есть то, ради чего стоит жить!» Другой его коллега, Сидни Шумейкер, как-то заметил: «Если, когда я пью вино, все сводится к информации о его химических свойствах, почему бы мне просто не прочитать этикетку на бутылке?»<sup>1</sup>. На эти вопросы у Деннета, как всегда, есть два типа эволюционного ответа: с одной стороны, говорит он, у человека имеются встроенные природой детекторы или склонности стремиться к одним видам вкусовых ощущений, цветовых оттенков, запахов, звуков и избегать других. С другой стороны, социум и его паттерны накладывают свой отпечаток на встроенные эволюцией диспозиции, культивируя подчас искусственные вкусы и удовольствия (дисгармоничная музыка, обжигающая рот пища и др.). Имея те же биологические корни, что и у наших предков, нынешние эмоции проявляются совсем по-другому. Были времена, когда рыцарь во имя любви к Прекрасной Даме, с которой он едва ли был знаком, с легкостью шел на смерть, чего трудно ожидать от современного человека. То же самое со вкусами: когда мы наслаждаемся хорошим вином, нас, конечно, интересует не информация о его химическом составе, «но [желание] *быть информированным* о его химическом содержании нашим из-

<sup>1</sup> Ibid. P. 383.

185

любленным способом. И наша преференция *в конечном счете* основана на склонностях, которые все еще встроены в нашу нервную систему, хотя их экологическая значимость могла быть утеряна давным-давно»<sup>1</sup>. А коли лежащие в основе эмоций биологические и социальные диспозиции не есть проявление внутреннего трансцендентного Я, они в принципе поддаются объективному гетерофеноменологическому анализу и воспроизведению; следовательно, мы можем помыслить робота, наделенного теми или иными эмоциональными предпочтениями.

### 3. Свобода воли и детерминизм

Очень много путаного, считает Деннет, философы наговорили о «свободе воли». Являются ли и мысли и действия личности результатами ее внутренних решений, актами свободного выбора, или они детерминированы цепью внешних причин? По этому поводу идет давний спор между «либертарианцами», защищающими свободную волю, «детерминистами», подводящими под нее различные внешние основания, и «компатибилистами», пытающимися совместить и то и другое. В

книгах «Пространство для свободы», «Головоломки» (Ch. 15), «Свобода развивается» Деннет подчеркивает, что и понятие «свобода воли», и понятие «детерминизм» соединили в себе множество часто несовместимых друг с другом смыслов, обросли толстыми слоями мифологии, не разобрав которые, невозможно выделить значения, с которыми можно работать. Живучесть понятия «свободная воля» объясняется не столько силой представленных в его пользу аргументов, сколько психологической потребностью самоутверждения: человеку комфортнее представлять себя хозяином своих решений и действий. Детерминизм тоже не без психологической подкладки — причины неудач легче переносить на внешние силы. По-

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 384.

186

этому его часто рисуют в виде различных пугал — фатализма, рока, неотвратимости и т.п. Деннет — физикалист и, по определению, детерминист. Но только в том общем смысле, что не допускает ничего «добавочного» к Вселенной — будь то ментальное сознание, квалиа или свободная воля. В то же время для него неприемлем детерминизм в его физическом, биологическом, экологическом, социальном варианте. Его Модель Множественных Набросков рисует чрезвычайно сложную картину «машины мозга», производящей сознание, в том числе осознанный выбор той или иной поведенческой стратегии.

Кем бы мы себя ни считали, свободными или несвободными, говорит он, мы являемся информационно-процессуальными системами, включенными в информационные системы природы, языка, социума и взаимодействующими с ними. Системы мозга представляют собой организацию разного рода включателей, выключателей и переключателей, работающих по принципу усилителя. Относительно малые, часто незаметные причины «на входе» могут производить относительно больше эффекта «на выходе» и модулировать активности, количественно превосходящие изначально полученную энергию. Полученная «на входе» информация может быть незначительна, и ее роль «толчка» не видна. Наблюдения только за «выходом» создают иллюзию, что полученный информационный результат — акт решения или выбора — не имеет причин и получен благодаря спонтанной самоактивности самости.

Мы не в состоянии проследить все задействованные в мозгу информационные процессы, непрерывно взаимодействующие с информационными процессами природы и культуры. Мозг постоянно генерирует информационные сети и, «ощупывая» изменения окружающей среды и собственного тела, ревизует имеющееся у него «знание». Еще Дарвин заметил, говорит Деннет, что природе свойственна «бессознательная селекция». Мозг тоже проделывает ее: какие-то внешние влияния не оставляют в нем следа, какие-то откладываются в долгосрочной памяти, а какие-то в вербальной или невербальной фор-

187

ме подсказывают решение. Индивиду они кажутся спонтанными и «свободными». Можно сказать, что человек постоянно занят «само-дизайном» и его совершенствованием, целью которого является адаптация к среде.

После выхода в 1995 году книги «Опасная идея Дарвина» на Деннета посыпались упреки в том, что он гипертрофирует адапционизм, недооценивает роль случайностей, что от Витгенштейна и Райла он дрейфует к биологическому детерминизму. Стивен Гулд приписал Деннету генетический детерминизм, нивелирующий роль воли, образования, культуры. Деннет счел эти упреки несправедливыми. В 2003 году он опубликовал книгу «Свобода развивается» («Freedom Evolves»), в которой дал развернутое разъяснение своему пониманию свободы и детерминизма<sup>1</sup>.

Деннет говорит следующее: «Существует Природа и существует Обучение». (There's Nature and there's Nurture). Вряд ли можно оспаривать, что все не детерминированное нашими генами должно быть детерминировано нашими окружающими условиями. Есть ли еще некий «X», добавка к тому, что мы есть в рамках детерминантов природной, социальной и культурной среды? Такой «X», считает он, есть. Это «Случай. Удача. Данный добавочный ингредиент важен, но его источник не нужно искать в квантовых внутренностях наших атомов или в воздействии каких-то дальних звезд. Он находится вокруг нас, в нашем шумном мире, где случайности игры в "орла" и "решку"

автоматически заполняют провалы спецификации, которые остались не зафиксированными нашими генами, а также провалы, не зафиксированные заметными причинами окружающих нас условий. Это особенно видно на способе оформления триллионов связей клеток нашего мозга»<sup>2</sup>. Учеными признано, что человеческий геном, будучи огром-

<sup>1</sup> К сожалению, в нашем распоряжении этой книги нет, и мы пользуемся материалами, которые он поместил на своем сайте в сети Интернет.

<sup>2</sup> Dennett D. The Mythical Threat of Genetic Determinism. Internet: <http://chronicle.com/free/v49/i21/21b00701.htm> . P. 2.

188

ным, в то же время слишком мал для спецификации в его генных рецептах всех сформировавшихся между нейронами связей. Происходит следующее. Гены специфицируют процессы, которые вызывают огромный популяционный взрыв нейронов, производя их во много раз больше, чем мозг в состоянии будет использовать; и эти нейроны выпускают наугад (псевдонаугад, конечно) их исследовательские щупальца. Многие из них вступают в связь с другими нейронами таким образом, что это оказывается заметно полезным (удаляются излишества). То есть Деннет хочет сказать, что природа (в том числе мозг) очень многое делает с излишествами, наугад, случайными путями. Но все это формы проявления ее алгоритма: «Исследование и Развитие» (Research & Development, сокращенно R&D).

На этом процесс поиска решения через случайности и отбор не кончается. Есть много другого в человеческом бытии, что зависит от случая. Прежде всего язык, обучение в определенных социальных условиях и существующая в культуре техника индоктринации. Все это так. Но когда и Nature и Nurture проделают всю свою детерминирующую работу и поделят сферы влияния, «останется ли после этого еще что-то, что есть Я?», задает вопрос Деннет. Останется, отвечает он. Это работа человеческого воображения на стыке Nature и Nurture, детерминантов и случайностей, или, в его терминологии, на основе «пульсации интуиций» (Intuition pumps). В его толковании «пульсация интуиций» не есть традиционная интуиция или интроспекция (этот смысл он отвергает). Скорее, это возникающая изнутри потребность выразить себя определенным образом, обозначить свое Я. Об этой потребности любят говорить люди творческих профессий, но она в той или иной мере присуща всем людям. На ее основе образуются «центры нарративной гравитации». В своем собственном творчестве и «мысленных экспериментах» Деннет видит результаты «пульсаций интуиции». Такие интуиции не являются машиной открытия, а скорее чем-то вроде средства самоубеждения или способа заставить людей видеть вещи вашими глазами, как только вы

189

увидели истину. «Рефлексия об истории философии показывает, — говорит он, — что величайшие "пульсации интуиции" на всем ее протяжении служили ее движущими силами»<sup>1</sup>. Будучи дотеоретическими, складывающимися на основе всего предшествующего опыта, они имеют в нем разветвленные причины. Редуцировать их к деталям этого опыта, конечно, невозможно. Они уникальны для данного человека.

#### 4. Эволюционные алгоритмы и творчество

Есть еще одно недоумение, которое вызывает у критиков деннетовская концепция самости: *как объяснить творчество человека?* Общепринято считать, что творчество, точнее самотворчество, является отличительной чертой homo sapiens. Из своей головы, своими руками человек создает новизну — артефакты, из которых состоит вся культура и цивилизация. Невозможно представить, чтобы сама природа или какой-либо другой биологический вид могли создать что-то подобное произведениям Шекспира.

Деннет говорит, что феномен творческой активности человека — одно из самых распространенных возражений против универсальности дарвиновской парадигмы. При общем натуралистическом настрое современных философов далеко не все готовы пойти до конца и принять радикализм вытекающих из дарвиновской парадигмы выводов. Когда Дарвин



опубликовал «Происхождение видов», книга была встречена в штыки креационистами, возмущившимися тем, что Дарвин заменил Замысел Творца ничем не ведающими слепыми процессами и поставил уникальную творческую активность человека в один ряд с простыми механизмами грубой природы. Креационис-

<sup>1</sup> Dennett D. *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge MA, 1984. P. 17. 190

ты, говорит он, верно схватили суть: на место Абсолютной Мудрости Дарвин поставил Абсолютное Неведение. Его гениальность состояла в том, что он увидел и доказал: Природа работает благодаря своим собственным механизмам — слепо, на ощупь и наугад, а ее алгоритмы не требуют ни Автора, ни Цели.

Если вывод о самодостаточном развитии природы, хотя и не без сопротивления, был принят, второй вывод, а именно, что сознание и все творчество интеллекта не есть исключение из правил природы, вызывал и вызывает большое сопротивление. И во времена Дарвина, и сегодня людям трудно смириться с мыслью, что плоды их чувств и воображения в конечном счете — игра нейронов, молекул или атомов. Сама идея, что все результаты человеческого гения могут быть, *в конце концов*, поняты как механически производимые продукты каскада алгоритмов генерируй-и-тестируй, вызывает искреннее возмущение многих людей, по другим вопросам достаточно глубоко и открыто мыслящих<sup>1</sup>. Как и креационистам, многим кажется кощунственной мысль о соединении «А» (artificial) и «I» (intellect), то есть чего-то искусственного с естественным. Например Дж. Сёрль, вполне рационально и научно мыслящий человек, сопротивляется идее «сильного AI», не может смириться с мыслью, что различие между механизмом, лишенным сознания, и человеком, наделенным сознанием, есть различие в дизайне (или программе).

Никто не спорит, что сознание представляет собой механизм поразительной сложности, но все же он не таков, чтобы его невозможно было помыслить без обращения к чуду или *elan vitae*, или вообще дистанцировать от компьютера. И сознание и компьютер находятся в пространстве дизайна Природы и действуют в соответствии с ее основным вектором или алгоритмом — *Исследование и Развитие* (Research & Development). Конечно, сегодняшние модели AI являются

<sup>1</sup> Деннет Д. По следу Дарвина. Где есмь Я? См. перевод в Приложении 1 данной книги.

191

ограниченными и упрощенными, многое еще неизвестно, возможны ошибочные стратегии в поиске оптимального дизайна компьютеров. Однако Сёрль говорит не о деталях и не об ошибочных стратегиях; он говорит о принципиальной невозможности сильного AI независимо от будущих исследований. И Сёрль и другие люди, сопротивляющиеся тому, чтобы отдать человеческий интеллект на откуп алгоритмам «Машины Дарвина», говорит Деннет, в сущности, рассуждают просто: поскольку ни один горшок не создал горшечника, ни одна машина не создала машиниста, значит, должна существовать какая-то превосходящая человека сила или надприродное свойство, превращающее материальную машину в мыслящего человека. Без этого дополнительного свойства телесная машина, полагают они, не в состоянии стать сознательной. На самом деле принятие дарвиновской парадигмы диктует другой ход мысли: отказ от каких-либо дополнительных сверхприродных свойств и признание того, что вся работа — и в природе, и в человеческом сознании, и в компьютере — осуществляется слепыми, не имеющими сознания механизмами.

Никто не спорит, что только человек творит артефакты. Тем не менее используемая при их создании поисковая деятельность не выходит за рамки алгоритма «Машины Дарвина» — «Исследование и Развитие» (R&D). «Если дарвиновские механизмы могут объяснить существование жаворонка во всей его изумительности, они, конечно, могут объяснить и существование оды в честь соловья. Поэма — замечательная вещь, хотя на самом деле не более удивительная, чем живой, поющий жаворонок»<sup>1</sup>.

Деннет убежден, что, несмотря на сопротивление идее сильного AI и эксцессы противников дарвинизма, люди рано или поздно придут к осознанию, что «дарвиновская теория творческого потенциала содержит в себе не только общее решение, она является единственным решением

проблемы, ко-

<sup>1</sup> Деннет Д. По следу Дарвина. Где есмь Я? С. 402—403.

192

торая является проблемой каждого из нас: *как может совокупное устройство тысяч миллиардов нейронов создать творческий ум, мое Я*<sup>1</sup>. Более того, дарвиновская парадигма содержит все, чтобы принять и творчество человека, и AI в пространство своего дизайна, ибо ее принцип «R&D» свойствен и машине.

Выступая в 2000 году с президентским посланием Американской философской ассоциации, заседание которой было посвящено теме творчества, Деннет привел следующий пример в пользу универсальности дарвиновской парадигмы, весьма показательный для вектора его аргументации. Хотя его изложение требует много места, приведем его полностью.

«Сегодня компьютеры еще не создали достойные внимания драматические произведения, однако уже есть искусственный игрок в шахматы — IBM Deep Blue — и искусственный композитор EM1, созданный Дэвидом Коулом. У них имеются результаты, в каких-то отношениях равноценные тем, которые достигнуты лучшими человеческими умами.

Кто победил Каспарова — чемпиона мира по шахматам? Его победил не Мюррей Кэмпбелл или кто-либо из его команды в IBM. Каспарова победил Deep Blue. Придуманные им шахматные игры оказались лучше, нежели дизайны игры, придуманные для него кем-либо из создавшей его команды специалистов. Ни один из них не может быть автором победной игры против Каспарова. А Deep Blue может. Все это так. Тем не менее я уверен, что когда Deep Blue одерживает победу над Каспаровым в шахматах, многие из вас, возможно, возразят, что грубые поисковые методы компьютера *совершенно* не похожи на исследовательские процессы, используемые Каспаровым в его прикидках возможных шахматных шагов. Однако это просто не так — или, по меньшей мере, это не так в единственном смысле, который мог бы иметь значение для контекста данного спора об универсальности дарвиновской перспективы применительно к творчеству. Мозг Каспа-

<sup>1</sup> Там же. С. 404.

193

рова состоит из органических материалов, и его архитектура в каких-то весьма важных аспектах непохожа на архитектуру Deep Blue, тем не менее, насколько я это знаю, мозг Каспарова представляет собой огромную параллельно работающую машину, которая в течение длительного времени выстроила превосходный вектор экономной поисковой техники, позволяющей не тратить время на ненужные ходы. Нет сомнения, что инвестиции в «R&D» имеют различные профили в этих двух случаях. Каспаров использует методы, заимствованные им из принципов прошлых игр с хорошим дизайном, поэтому он в состоянии распознавать и пользоваться большим запасом знания, с тем чтобы проигнорировать огромную часть игрового пространства, через которое Deep Blue еще нужно терпеливо продираться *seriatim*. В форму поиска, который ведет Каспаров, свои драматические изменения вносят "озарения", тем не менее это обстоятельство не образует *"совершенно иное"* средство творческого процесса. Всякий раз, когда исчерпывающие поиски Deep Blue закрывают тот тип пути, для определения которого у него имеются средства распознавания (трудная, но не невозможная задача), когда это возможно, он может заново использовать «R&D», то есть поступать так же, как поступает Каспаров. Многие из аналитической работы Deep Blue проделано для него его дизайнерами и придано ему в виде внутреннего багажа, однако точно так же Каспаров очень многое заимствовал из сотен тысяч часов личного времени, потраченного его предшественниками на исследование шахмат, и из переданного ему игроками в шахматы, тренерами и книгами. В этой связи заслуживает внимание высказанная недавно Бобби Фишером идея. Он предлагает воссоздать игру в шахматы в изначальной чистоте (информационной — *Н. Ю.*), но с условием, чтобы в начале каждой игры главные фигуры *случайно* располагались в заднем ряду (случайно, но в зеркальном отражении для черных и белых фигур). Изменение условий как для людей, так и машин тотчас же сделает массив запечатленных в памяти шахматных открытий почти совершенно излишним, поскольку только в редких случаях ка-

194

кое-либо из этих профессиональных знаний (lore) будет востребовано в игре. Один будет отброшен назад и будет искать для себя опору в фундаментальных принципах, другому потребуется проделать значительно большее количество дизайнерской работы в реальное время — до сигнала часов. И вовсе неясно, кому это изменение в правилах игры даст больше преимуществ — человеческим существам или компьютерам. Все зависит от того, какой из этих типов игроков в шахматы в большей мере опирается на то, что в итоге есть механическая память — опора с минимальным пониманием «R&D» предшествовавших им искателей.

Фактом является то, что поисковое пространство для шахмат слишком велико, для того чтобы даже Deep Blue мог исследовать его исчерпывающе в реальный отрезок времени, поэтому, как и Каспаров, он обрезает какие-то поисковые пути и прибегает к риску калькуляции и, подобно Каспарову, он часто прибегает к этим рискам в до-вычислительном виде. Предположительно оба они прибегают к большому количеству расчетов "грубой силы" на основе их очень разных архитектур. В конце концов, что эти самые нейроны знают о шахматах? Любая совершаемая ими деятельность должна быть тем или иным видом работы грубой силы»<sup>1</sup>.

## 5. Где находится мое Я?

### Телепортация и бессмертие

Если согласиться с тем, что самость лучше всего рассматривать как активность виртуальной машины или «мягкую программу» органического компьютера — мозга, возникает одна малоприятная перспектива для нашего самообраза. Деннет выразил ее в остроумном слогане: «Ученый человек — это только способ, каким одна библиотека создает другую библио-

<sup>1</sup> Деннет Д. По следу Дарвина. Где есмь Я? С. 8.

195

теку»<sup>1</sup>. Ему самому, признается он, психологически не очень комфортно думать, что «его мозг является разновидностью кучи навоза, в котором личинки идей других людей обновляют себя, прежде чем посылать свои собственные копии в информационную Диаспору»<sup>2</sup>. Получается что-то вроде обворовывания сознания и моей значимости как автора и критика. Слоган Деннета мы бы представили в более пугающем виде: «Software», составляющий самость, — это только способ, каким одна компьютерная программа создает другую компьютерную программу, та — третью и так до бесконечности. Не говоря уже о другой, моральной перспективе: приписав самость компьютеру, мы в соответствии с существующими в обществе нормами должны признать за ним моральные права, а вместе с ними и ответственность. (См. об этом подробнее в гл. VIII).

Деннет признает, что у него нет простого рецепта для снятия этих сомнений. Хотя серьезных контраргументов против идеи телепортации он тоже не видит. И вот почему. Один из вопросов, который возникает у людей по поводу нарративной трактовки самости, таков: «Где находится такая вещь, как моя *собственная нарративная саморепрезентация?*» «Нигде» или там, где вы хотите, отвечает он. Примечателен в связи с этим его диалог с «Отто» — персонажем, введенным им в книгу «Сознание объясненное» в качестве выступающего от имени здравого смысла и фолк-психологии оппонента.

Отто говорит:

«Беда с твоими центрами гравитации состоит в том, что они не являются реальными; они — фикции теоретика».

На что следует такой ответ Деннета:

«Это не беда центров гравитации; это их преимущество. Они являются потрясающими фикциями, которыми каждый мог бы гордиться. Даже воображаемые персонажи литературы не столь удивительны. "Называй меня Дэном", ты слы-

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 202.

<sup>2</sup> *Ibidem*.

196

шишь из моих уст при нашем знакомстве, и ты повинешься, но не тому, что мои губы являются Дэном или мое тело является Дэном, но фикции теоретика, сотворенной не только мной, а моим мозгом, взаимодействующим в течение многих лет с моими родителями и друзьями».

Отто возражает:

«Все это прекрасно для тебя, но я — совершенно реален. Возможно, что я сотворен социальными процессами, на которые ты как раз ссылаешься (я должен быть, если я не существовал до моего рождения), но нечто, сотворенное этим процессом, есть реальная самость, а не просто придуманный характер».

На самом деле, разъясняет ему Деннет, тебя волнует не реальность самости, а реальность метафизической основы моральной ответственности. Людям представляется, что без самости или души рушатся все устои морального сообщества. Между тем как задача конституирования самости, принимающей на себя ответственность, относится к сфере социальных и образовательных проектов. К ней следует подходить серьезно и прагматически, не затуманивая голову противоречивой метафизикой особой, отдельно существующей души.

Но как тогда быть с телом? По Деннету, являясь виртуальными машинами, наши самости, как и физические центры гравитации, не обязательно должны быть привязаны к какому-либо конкретному физическому медиуму.

Отто недоумевает:

«Получается, что я не существую?»

Ответ Деннета примечателен.

«Конечно, ты существуешь, — успокаивает он Отто, — ты сидишь передо мной в кресле, читаешь мою книгу и возражаешь мне. Более того, в случае твоего согласия с моим взглядом на самость на как виртуальную реальность, перед тобой, в принципе, открывается возможность телепортации и бессмертия. Твое нынешнее телесное воплощение хотя и является необходимым предварительным условием существования, не является необходимым условием неопределенно долгого суще-

197

ствования. Впрочем, если ты считаешь себя душой — жемчужиной нематериальной субстанции, тогда можешь надеяться на бессмертие в виде *virtus dormitiva* духовного вещества. Если же ты считаешь себя жемчужиной материальной субстанции — группой атомов мозга, твоя смертность зависела бы от чисто физических сил. Но вот если ты будешь думать о себе как о центре нарративной гравитации, твое существование будет поставлено в зависимость от наличия устойчивого нарратива (скорее, как «Тысяча и одна ночь» в виде одной сказки), который мог бы теоретически пережить бесконечное множество смен медиумов и быть подвергнут телепортации». «Если то, что ты есть — организация информации, структурировавшая контрольную систему твоего тела (или, формулируя мысль в более провокативной форме, если то, что ты есть — программа, которая продолжает управлять твоим компьютером мозга), тогда ты, в принципе, мог бы пережить смерть своего тела столь же неповрежденным, как и программа, которая может сохраняться после разрушения компьютера, на котором она впервые была создана и запущена»<sup>1</sup>. Некоторые ученые, в частности Роджер Пенроуз (1989), замечает Деннет, нашли мой взгляд на телепортацию контринтуитивным и «пугающим». Однако, если заводить разговор о потенциальном бессмертии, которого, надо полагать, все страстно желают, альтернативные взгляды просто не могут быть защищены.

С нашей точки зрения, недоумение Пенроуза можно частично рассеять, если посмотреть на идею программного бессмертия и телепортации Деннета под углом зрения историка философии. Историк это напоминает неоднократно проигрывавшуюся схему платонизма. На вопрос Отто, где находится его самость, Деннет отвечает: «Нигде». На самом деле у него есть более содержательные термины для обозначения ее местонахождения — «Информационная Диаспора» или «Мемосфера». Телесная составляющая сознания — процессы мозга — имеет пространственно-временную локализацию, содержатель-

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 430.

198

ная — когнитивная составляющая сознания — вне пространства и вне времени. Не имея физических признаков, эта составляющая тем не менее тоже представляет собой процесс — операцию с информационными единицами, т.е. меме. Процесс этот отнюдь не хаотичен. При всем том, что сферу воображаемого или виртуального отличает неопределенность (здесь не действует принцип двусмысленности, то есть либо факт есть, либо его нет), она не является иррациональной: ей присуща истинностно-сохраняющая функция, в ней возможны интерпретация и экстраполяция с использованием законов логики, здесь существует своя иерархия объектов (выводимые сущности физики — *illata* — имеют более высокий статус, нежели просто воображаемые сущности — *abstracta*).

Судя по работам Деннета, он внимательно читал Платона (многие его образы в изображении сознания, например образ идей-птиц, он находит великолепными). Если присмотреться внимательнее, не обращая внимание на современный жаргон, можно обнаружить признаки родства между «Информационной Диаспорой» или, как он ее еще называет, «Мемосферой» и идеальным царством форм Платона. Напрашивается также аналогия позиции Деннета с позициями других платоников (или скрытых гегельянцев), выносивших идеальное содержание сознания на особый уровень бытия, развивающийся по своим законам, например, с Дж. Сантаяной и К. Поппером. К слову сказать, Деннет обрушился с резкой критикой на метафизику «Трех миров» Поппера. Однако предметом неприятия является не столько Мир 3, сколько Мир 2 — то есть мир ментальных психических состояний (феноменов), выполняющий посредническую роль между Миром 1 и Миром 3.

Более подробно мы будем об этом говорить в разделе «К. Поппер и Д. Деннет: "архитектура сознания" согласно "Открытой" и "Закрытой" Вселенной».

## VI. Интенциональность как установка рациональности

Интенциональность — более фундаментальный феномен, нежели сознание.

*Дэниел Деннет*

Приверженность доктрине первичности внутренней интенциональности является главным источником путаницы в современной англоязычной философии сознания.

*Дэниел Деннет*

### 1. Бихевиоризм и интенциональность

В предыдущих главах мы рассказали о деконструктивистском проекте Деннета по разрушению менталистской онтологии и феноменологии и низвержению самодостаточной самости с пьедестала главного героя западной философской традиции. Были обрисованы основные блоки его конструктивного проекта объяснения сознания — гетерофеноменология, функционализм, Модель Множественных набросков, аналогия естественного и искусственного интеллекта, динамика процесса сознания в свете «Машины Дарвина» и «Машины Тьюринга». В нашем рассказе о конструктивной стороне проекта отсутствует важный блок, выполняющий связующую роль в понимании содержательного отношения сознания с миром. Речь идет о концепции интенциональности. По логике этапов

200

работы Деннета и по фундаментальности этой концепции для всей его философии, эта тема должна была фигурировать в нашем тексте в самом начале, поскольку речь идет о главном — о содержательности и семантической наполняемости сознания. У нас она идет позднее, что объясняется интересом к метафизике его проекта (обоснование единства мира), а она лучше просматривается на основе представленного ранее материала.

По признанию комментаторов, концепция интенциональности является самым оригинальным вкладом Деннета в философию. Многие — не только философы — считают, что она имеет прикладную методологическую значимость и применима в самых различных поведенческих областях — социологии, психологии, когнитивной этологии, теории игр, психиатрии и др. В пределах одной главы мы лишены возможности останавливаться на всех разбираемых им



вопросах и ограничимся только несколькими. Их выбор продиктован задачей показать в действии используемые им концептуальные методы и средства, в частности эволюционный когнитивизм и дуальную перспективу, в создании недихотомичной картины человека, природы и культуры. В рамках этой задачи мы рассмотрим его теорию «установок», аргументы в пользу объективации интенциональности, трактовку интенциональной установки как позиции интерпретивизма и одновременно как «свободно плавающей рациональности». Коснемся особенностей эволюционного подхода к интенциональности, а также понимания «первичной» и «производной» интенциональности. В конце главы приведем ряд возражений критиков эволюционного аргумента Деннета.

В словарях интенциональность обычно определяется как направленность на что-то или кого-то. В речи она выражается в фразах «я верю, что айсберги существуют», «я намерен произвести ремонт», «я люблю кататься на коньках», «он вызывает у меня симпатию» и т.п. В отличие от фраз «айсберги плавают», «на коньках катаются», «некоторые люди вызывают симпатию», первый набор фраз представляет собой

201

предложения о чем-то, к чему мы высказываем свое отношение. Эту разницу люди заметили давно. На ее философско-теоретический смысл впервые обратил внимание Франц Brentano (1874). Brentano подошел к сознанию с другой, нежели Декарт, стороны и увидел дополнительный ракурс. Он указал на то, что всем ментальным феноменам свойственна интенциональность, у физических вещей она отсутствует. В «*направленности на*» он увидел признак онтологического существования ментального. Понятие «интенциональность» у него имело несколько смыслов. Во-первых, объектами интенциональности могут выступать любые сущности, независимо от их природы: ими могут быть как физические вещи, так и содержание сознания или значения; во-вторых, в силу интенциональности ментальное не может быть сведено к физическому. Последний тезис традиционно толкуется как тезис *нередуцируемости* («тезис Brentano»). Из него логически следовало, что Вселенная раздвоена на ментальное и физическое, а это значит, что знание расколото на «науки о духе» и «науки о физическом» по онтологическому признаку.

Обобщая, можно сказать, что на «тезис Brentano», или, как его еще называют, «боевого коня» философии сознания, мысль XX века отреагировала двумя интеллектуальными движениями: феноменологией и бихевиоризмом. Феноменология стала разрабатывать субъективный, внутренний, ментальный аспект интенциональности, бихевиоризм взял курс на преодоление раскола мира на физическое и не-физическое с помощью различных объективистских стратегий. Первое движение было характерно, прежде всего, для европейской мысли, второе — для англоязычной.

Бихевиорист Беррес Скиннер сконцентрировал внимание на возможностях объективного анализа психологического поведения. Его посылка состояла в том, что описание психологических актов должно ограничиваться внешне наблюдаемыми факторами без обращения к внутреннему (иллюзорному) ментальному агенту. Эти факторы он свел к получаемым извне

202

стимулам и реакциям на них организма, т.е. к чисто механическим операциям и регулярностям в действиях и поведении субъектов. В его операционалистской схеме понятие «интенциональность» было ненужным, и он отбросил его вместе с понятием «ментальное». В логическом бихевиоризме Джилберта Райла в нем тоже не было надобности. Уилард Куайн, на том основании, что интенциональные идиомы (я хочу, я радуюсь чему-то, я сокрушаюсь о чем-то) проявляют «странности», не поддаются строгому логическому анализу и не могут быть встроены в рамки будущего (идеального) нейрофизиологического языка науки, считал, что они должны быть отброшены, а феномены, которые они якобы описывают, признаны химерическими<sup>1</sup>.

Далеко не все аналитики сочли «интенциональность» неработающим понятием. В развернувшихся вокруг идей Л. Витгенштейна и Дж. Райла дискуссиях некоторые из них обратились к анализу интенциональных идиом и попытались дать свое, отличное от Brentano, определение (Р. Чишолм, П. Гич, Э. Энском). В аналитической философии постановка проблемы интенциональности, как,

впрочем, и всех других, имеет свою специфику, обусловленную тем, что она переводится с онтологического уровня — разговора о природе физических и ментальных объектов — на языковый уровень. Ставятся, например, такие вопросы: каков семантический смысл высказываний с интенциональными терминами «верю, что», «желаю, чтобы», «забочусь о чем-то», «люблю кого-то» и высказываний, в которых они отсутствуют, например, «снег падает», «деревья растут»? Возможно ли предложения типа

<sup>1</sup> Куайн писал: «Можно принять тезис Brentano или в качестве демонстрации неустраимости интенциональных идиом и важности самостоятельной науки, которая бы занималась интенциональностью, или в качестве демонстрации бесосновательности интенциональных идиом и бессодержательности такой науки. Я, в отличие от Brentano, разделяю второе отношение». (Куайн У. Слово и объект. М., 2000. С. 251).

203

«ненавижу что-то», «стремлюсь к чему-то» без изменения значений перевести в любые предложения, доступные для интерсубъективного наблюдения и оценки? Что следует считать референтами интенционального языка, если сравнивать его с языком физики?

Языковый и логический уровень обсуждения понятия «интенциональность», прояснив и уточнив смыслы его употребления, все же не снял зафиксированную Brentano дихотомию интенционального и физического. Она проявилась в концентрации внимания на дилемме: следует ли считать интенциональность внутренним свойством индивидуальной личности, или ее можно рассматривать как нечто внешнее, что может исследоваться объективными интерсубъективными методами (с позиции третьего лица)? Для лингвистически ориентированной аналитической философии решение этой дилеммы принципиально значимо еще и потому, что она имеет отношение к источнику содержательности и значений. В зависимости от решения дилеммы «внутреннее—внешнее» (intrinsic—extrinsic), философы, занимающиеся сознанием, разделились на два лагеря: считающих, что интенциональность является внутренним (intrinsic), изначально присущим сознанию свойством (Дж. Фодор, Дж. Сёрль, С. Крипке, Т. Нагель и др.), и тех, кто считает, что интенциональность есть внешняя (extrinsic) наша установка (У. Куайн, У. Селларс, П. Черчленд, Р. Рорти, Д. Дэвидсон, Дж. Хогелэнд, Р. Милликэн, М. Минский, Д. Хофстадтер, Д. Деннет и др.). Деннет, конечно, находится во второй группе.

Во многом восприняв объективизм, натурализм и физикализм Куайна, Деннет тем не менее категорически не согласился с ним в том, что логические трудности определения интенциональных идиом делают их неработоспособными. Жесткий физикализм закрывает глаза на специфический факт— целенаправленность человеческого поведения. Люди, желающие идти по пути объективизма, не должны уповать только на прогресс нейрофизиологии и отмахиваться от тезиса Brentano; задача отыскания убедительных теоре-

204

тических средств для наведения мостов через обозначенную им пропасть между миром ментального и миром физического остается актуальной. Более того, одно из твердых его убеждений, принципиальных для всей его философии, состоит в том, что рассмотрение интеллектуальной и целенаправленной деятельности через призму интенциональности является самым плодотворным подходом, дающим возможность достичь объективное понимание, недостижимое никакими другими способами. Поэтому в философии сознания этому понятию должна быть предоставлена ключевая роль. Наиболее важные для себя подсказки он нашел в работах Родерика Чисолма, Чарльза Тейлора и Элизабет Энском<sup>1</sup>. Хотя уже в ранней работе он сформулировал свой собственный подход к интенциональности, существенно отличающийся от подходов, представленных этими авторами. Он остался верен ему и в книге «Интенциональная установка».

«Узурпировав», по его словам, термин Ф. Brentano и найдя неудовлетворительным данное им (и во всей традиции Э. Гуссерля и феноменологии) толкование интенциональности, он предложил применить к исследованию интенциональности новые инструменты — дарвиновский эволюционизм и идеи функционализма машины Тьюринга. Цель такой стратегии многоплановая: связать ментальное и физическое, науки о физическом и «науки о духе», доказать отсутствие

пропасти между фактами природы и артефактами, естественным и искусственным интеллектом; одним словом, обосновать континуальность мира и единство знания. Не говоря уже о том, что с помощью понятия «интенциональность» им решаются эпистемологические задачи обоснования рациональности.

<sup>1</sup> Витгенштейн Л. Философские исследования. М., 1999; Райл Г. Понятие сознания. М., 2000; Chisholm R. Perceiving. Philosophical Study. № 9. 1957; Anscombe G.E.M. Intention. Oxford, 1963; Taylor Ch. The Explanation of Behavior. L., 1964; Geach P.T. Intentional Identity // Journal of Philosophy. LXIV. 1967.

205

В работах Деннета 60—70-х годов об интенциональности было множество нестыковок, что тут же было отмечено критиками. В последующих книгах — «Интенциональная установка» (1987), «Сознание объясненное» (1991), «Опасная идея Дарвина: эволюция и смысл жизни» (1995), в материалах, фигурирующих в Интернете, он расширил и детализировал тему интенциональности<sup>1</sup>.

## 2. Физическая установка, установка дизайна и интенциональная установка

Разговор о концепции интенциональности Деннета лучше всего начать с его трактовки «установок» — исходном моменте его варианта бихевиоризма. Занимаясь объяснением мира, говорит он, мы прежде всего занимаемся объяснением *поведения объектов*: нас интересует, как можно предсказать их поведение для установления контроля и использования их для своих нужд. Людям кажется, что в своих исследованиях они раскрывают сущность или природу объекта (атома, клетки, человека), на самом деле они исследуют его поведенческие проявления при тех или иных условиях.

В постижении мира в нашем распоряжении имеются три объяснительные стратегии или *установки* (stances)<sup>2</sup>, на основе которых мы предсказываем поведение объектов: *физическая установка, установка дизайна и интенциональная установка*. Все эти установки зависят от нашего *прагматического выбора и целей*.

<sup>1</sup> Dennett D. The Intentional Stance. L., 1987; Dennett D. Consciousness Explained. Boston, 1990; Dennett D. Darwin's Dangerous Idea. Evolution and the Meaning of Life. L., 1995.

<sup>2</sup> Слово «stance» в переводе с английского означает позицию по отношению к чему-то, положение тела, интеллектуальное или эмоциональное отношение к чему-то.

206

Можно принять физическую установку. Мы прибегаем к ней, когда произносим такие, например, фразы: «если нападает много снега, эта ветвь обломится», «если растение оставить без влаги, оно погибнет», «если соединим два атома водорода и один кислорода, получим воду». В этом случае объяснения являются каузальными, а предсказания строятся на основе сведений о действительных физических свойствах тех или иных объектов и эффективны в зависимости от применения имеющегося у нас *знания о законах природы*.

Когда нужно, исходя из знания структурной организации объекта, объяснить и предсказать его функциональное поведение, мы прибегаем к установке дизайнера<sup>1</sup>. Установка дизайнера позволяет делать различные предсказания, но все они схожи в том, что опираются на понятие *функции*, зависимой от цели и являющейся *телеологической*. Установка дизайнера чаще всего применяется при предсказании функционального пове-

<sup>1</sup> В русском языке слово «дизайн» используется чаще всего в эстетическом смысле — как композиция, художественное убранство или, реже, как инженерный проект. В английском языке слово design имеет много значений — замысел, план, намерение, проект, конструкция и др. Термин design прочно вошел в философский лексикон после публикации книги Герберта Саймона «The Sciences of the Artificial» (1969), в которой тот защищал тезис о необходимости радикальной переориентации всех наук с целью превращения их в «науки о дизайне», то есть науки об искусственно сконструированном. Этот тезис он обосновывал тем, что мир, в котором мы живем, в большей мере является артефактом, нежели природным миром. Идеи Саймона оказали большое влияние на Деннета, однако его не устроил прямолинейный социологизм Саймона. Соответственно термин «дизайн» (и «установка дизайнера») у Деннета имеет как инженерно-проектировочный, так и биологический смысл, точнее, относится к функциональной

организации, а не только к конструкциям. Он применим и к компьютеру, и к биологическим существам. Учитывая этот философский смысл термина design, мы используем его без перевода.

207

дения механических объектов: «как только каретка машинки подойдет к концу, должен зазвенеть звоночек», «если убрать из машины колесо, она не будет двигаться». Ее существенным моментом является то, что «предсказания делаются, исходя исключительно из знания или предположения относительно функционального дизайна системы, безотносительно к физической конституции или внутреннему устройству тех или иных объектов»<sup>1</sup>. Мы можем применять ее не только к механическим объектам, но и к биологическим и даже к таким большим системам, как эволюционная организация природы. В нашем распоряжении есть также *интенциональная установка* (intentional stance). Мы прибегаем к ней, когда не можем объяснить поведение объекта на основе первых двух установок. Это происходит тогда, когда имеет место *целенаправленное поведение*. Точнее, когда нам лучше и эффективнее объяснять и предсказывать поведение системы, допуская наличие в ней цели и способности достигать эту цель. Объяснения на основе интенциональной установки не являются каузальными, они представляют собой интерпретацию объекта. «Элемент интерпретации... должен распознаваться в любом использовании интенционального словаря»<sup>2</sup>. Поэтому у интенциональной установки нет такого онтологического статуса, какой имеют физическая установка и установка дизайна (вот почему она была отвергнута в бихевиоризме Скиннера и Куайна), однако без нее не обойтись в случае поведенческих систем. Самое существенное здесь то, что понятие «интенциональность» используется не как онтологический признак ментального, как это было у Brentano, не в смысле внутренне присущего агенту свойства, а в смысле оптимальной объективистской стратегии (лингвистической, гносеологической, прагматической) в интерпретации и предсказании поведения любой системы. «Какая-либо вещь является интенциональной

<sup>1</sup> Dennett D.C. *Brainstorms. Philosophical Essays on Mind and Psychology*. Cambridge MA, L., 1986. P. 4.

<sup>2</sup> Dennett D. *The Intentional Stance*. P. 342.

208

системой только в отношении к стратегии того, кто пытается объяснить и предсказать ее поведение»<sup>1</sup>. Таким образом, это «приписывающая» установка, взгляд «с позиции третьего лица», гетерофеноменология.

Интенциональная установка применима не к каждой вещи: мы не можем применить ее, например, к камню (здесь требуется физическая установка), но только к *интенциональным системам*. Под интенциональной системой Деннет понимает «систему, чье поведение может быть, по меньшей мере иногда, объяснено и предсказано, опираясь на приписывание системе верований и желаний (надежд, страхов, интенций, подозрений)»<sup>2</sup>. При этом не имеет значения, с кем или с чем мы имеем дело. Главное, чтобы это была поведенческая система. Мы можем приписывать в одном и том же теоретическом смысле верования и желания человеческим существам, собакам, насекомым, компьютерам и даже термостатам, поскольку их можно считать системами, поведение которых можно предсказать при помощи *интенциональной стратегии*, т.е. при помощи интерпретации их поведения как (более или менее) рационального следствия их верований, желаний и обстоятельств.

У «приписывающего» толкования понятия «интенциональная установка» Деннет видит много преимуществ. Прежде всего, простота. Вместо того чтобы беспокоиться по поводу тысяч каузальных взаимоотношений (физическая установка Куайна) или сотен правил и определителей, встроенных в процесс (установка дизайна), все, о чем следует беспокоиться — иметь скромное верование, что поведение данной системы согласуется с характерными для нее положением, правилами и условиями среды. Что она также имеет некоторую минимальную степень рациональности, или само-интереса, для достижения ее целей при данных обстоятельствах. Другое преимущество «приписывающей» интенциональной установ-

<sup>1</sup> Ibid. P. 3—4.

<sup>2</sup> Ibid. P. 3.

209

ки состоит в ее *нейтральности* в отношении метафизических и онтологических вопросов, а также вопросов, касающихся материалов, композиции, конституции, сознательности, морали или божественности сущностей, попадающих под эту категорию. Все эти характеристики, подчеркивает он, принимаются как *допущения*, в противном случае интенциональная установка не работает, поскольку нельзя привести четкие критерии для проверки атрибутов интенциональной установки.

Интенциональные системы могут быть разных степеней сложности и разных уровней. Самая сложная система — это, конечно, личность. Личность может рассматриваться и с точки зрения физической установки (когда вычисляется, например, нагрузка лифта и нам нужно обеспечить его безопасность), и исходя из установки дизайнера (когда проводится хирургическая операция и важно знать особенности функционирования органов), и как интенциональная система (когда личности предлагается принять участие в обсуждении какого-либо вопроса и нас интересуют ее мнение). Не исключено применение, в зависимости от контекста и целей, других установок, например моральной. Моральная установка — это подкласс интенциональных установок. Допускается применимость идеи интенциональной установки к социальным и экономическим институтам, объектам эволюционной биологии, куль-тур-антропологии и др.

Интенциональная установка вполне применима к поведению неживых систем, например к играющему в шахматы компьютеру. Будучи противником по игре и пытаясь предсказать его возможные шаги, можно принять по отношению к нему, по меньшей мере, три разные установки. Можно применить физическую установку, однако она ничего не дает; знание материалов, из которых сделаны компьютеры, или законов электрофизики не помогут предугадать его шаги в шахматной игре. Можно применить установку дизайнера. Если точно известен проект компьютера (в том числе его непостоянная часть — программа) и если у него нет поломок, можно предсказать его реакцию на любой шаг, совершаемый в соответствии с инструкциями про-

210

граммы. Установка дизайнера позволяет делать различные варианты предсказаний, но все они зависят от цели, ради которой создавался компьютер, и от функции, которую его конструкция должна реализовать. Опираясь на эту установку, предсказываются неполадки в его конструкции и производится ремонт. Однако в наши дни играющие в шахматы компьютеры стали весьма сложными, из-за чего, исходя из установки дизайнера или физической установки, их шаги не всегда могут предсказать даже их проектировщики. Тем не менее у нас есть шанс победить машину в шахматной игре в случае принятия по отношению к ней интенциональной установки, прикидывая ее возможные хитроумные шаги и, конечно, зная цели и правила игры в шахматы. Иначе говоря, когда нет надежды предсказать реакции машины на основе знания физики или системного устройства, поражения в игре с ней можно избежать, приняв ее за *интеллектуального оппонента*. Последнее означает допущение, что машина является агентом, способным к *рациональному поведению*.

В свете допущения, что машину целесообразно рассматривать в виде агента, которому можно приписывать рациональные мыслительные действия, компьютер вполне возможно включать в число интенциональных систем. Более того, «это означает, что информацию, которой «обладает» компьютер, можно назвать верованиями компьютера, а его цели и подцели — его *желаниями*»<sup>1</sup>. На первый взгляд, говорит Деннет, может показаться, что «обладание» в этом случае является слабым и невинным понятием информационного багажа; на самом деле оно является, и должно быть, «эпистемическим обладанием» — аналогичным верованию. Например, про француза, имеющего «Британскую энциклопедию», но не знающего английского языка, можно сказать, что он обладает «Британской энциклопедией». Однако если и существует такой смысл обладания, он не является достаточно сильным в сравнении с тем видом обладания, какой должен быть у компью-

<sup>1</sup> Dennett D. The Intentional Stance. P. 6.



211

тера в отношении *использования* информации, которую он «выбирает», проделывая шахматный ход. Поэтому «желания» целенаправленно работающего компьютера должны быть специфицированы интенционально. «Успех этой установки, конечно, есть дело, которое улаживается прагматически, без апелляции к тому, имеются ли *в реальности* у объекта верования, интенции и тому подобное; поэтому безотносительно к вопросу, может или не может компьютер быть осознающим или иметь мысли, или иметь желания, некоторые компьютеры, несомненно, являются интенциональными системами, поскольку они являются системами, чье поведение может быть предсказано, причем наиболее эффективно предсказано при принятии по отношению к ним интенциональной установки»<sup>1</sup>. Одна из излюбленных тем в диспутах — мыслит ли компьютер «реально» или «нереально». Многие философы выражают скепсис относительно его мышления. Признавая определенные интеллектуальные достоинства компьютера, они полагают, что его нельзя считать «реально» обладающим верованиями и желаниями и вообще мыслящим. Деннет считает, что данное им определение снимает эти сомнения: в определении не говорится, что интенциональным системам *реально* присущи верования и желания; в нем сказано только о том, что их поведение можно предсказывать и объяснять, *приписывая* им верования и желания. Одни и те же предсказания будут получены и тогда, когда мы выстраиваем их в терминах верований и желаний, и тогда, когда мы выстраиваем их в терминах информационного багажа компьютера и его специфических целей. Сегодня интенциональное объяснение и предсказание поведения играющих в шахматы компьютеров является общепринятым и практически работающим методом в науке. Оно применяется безотносительно к физической «субстанции» компьютеров, к их происхождению, к наличию или отсутствию у них моральной позиции, сознания или самосознания, к детерминированности или недетерминированности

<sup>1</sup> Dennett D.C. *Brainstorms*. P. 238.

212

их поведенческих операций. «Решение принять эту стратегию является чисто прагматическим: оно не является внутренне правильным или ложным»<sup>1</sup>.

Вознаваемые играющему в шахматы компьютеру почести, разумеется, не следует расценивать в том смысле, что он является двойником сознания, адекватной моделью интеллектуальной деятельности человека или психической жизни животного. Речь идет только о том, что, учитывая его способность выполнять интеллектуальную работу, по отношению к его поведению целесообразно принимать такую же интенциональную установку, какую мы принимаем по отношению к интеллектуальному поведению людей. «Порою чисто физическая система может быть столь сложной, и в то же время столь организованной, что для предсказания ее действий мы находим более удобным, лучше объясняемым и прагматически необходимым обращаться с ней таким образом, как если бы она имела верования и желания и была рациональной»<sup>2</sup>.

Достоинство расширительного смысла понятия «интенциональная система», как мы уже сказали, Деннет видит в его простоте и относительной *метафизической нейтральности*. Проще решать вопрос, может ли машина быть интенциональной системой, нежели метафизический вопрос, может ли она *реально* думать или быть морально ответственной. Прагматическая стратегия использования установок хороша еще и тем, что позволяет по своей воле и в зависимости от прагматических целей менять типы установок. Например, находясь в роли шахматного оппонента компьютера (прикидывая возможные ходы машины), можно использовать интенциональную установку; работая над совершенствованием проекта компьютера — применять установку дизайнера; выполняя роль ремонтника — физическую установку. Разумеется, что объясняемые и предсказываемые феномены, исходя из той или иной выбранной установки, не являются одинаковыми. Целенаправленные акты

<sup>1</sup> Ibid. P. 7.

<sup>2</sup> Ibid. P. 7—8.

213

человека, например, не есть то же самое, что физиологические отправления; и то и другое не то же

самое, что физическое действие. И конечно, использование разных установок требует разных словарей, методологий наблюдения и объяснения. Словарь, используемый в установке дизайна, не работает при применении интенциональной или физической установок. Особый вопрос о соотношении установок. По сути, это вопрос об отношении физической картины мира, знаний о структурно-функциональной организации объектов и знаний о целенаправленном (интенциональном) их поведении. Деннет убежден, что принятие методологии установок — единственно эффективный путь к снятию обозначенного Брентано раскола знания на естественные науки и науки о человеке.

### 3. Интенциональность, эволюционный аргумент и «свободно плавающая рациональность»

Итак, условием эпистемической и прагматической успешности интенциональной установки, по Деннету, является не просто приписывание поведенческой системе верований, целей и т.п.; главное состоит в *допущении ее рациональности*. Понятие «рациональность», как мы знаем, уже не одно десятилетие дебатруется в литературе — без большого успеха. Деннет предложил свой подход к пониманию рациональности, основанный на понятии «интенциональность». Рассмотрим его.

Как мы уже сказали, интенциональная установка, по Деннету, представляет собой прагматически полезную интеллектуальную стратегию предсказания и объяснения поведения любой интенциональной системы — живой или неживой. Он оговаривает, что использование интенционального языка включает в себя *интерпретацию*. Это означает, что наше приписывание интенциональности объекту не является фактуальным и дескриптивным, следовательно, здесь не работают критерии

214

истинности или ложности. Интерпретация, как известно, обусловлена социокультурными факторами и предполагает «социализированную», а не «натурализированную» эпистемологию. Деннета не устраивает крайний социологизм с его неизбежными спутниками — инструментализмом и релятивизмом. К инструментализму у Деннета осторожное отношение. В ранних работах он придерживался инструментализма, в дальнейшем он высказал сомнение в его плодотворности, ибо в радикальном виде тот открывает дорогу безудержному релятивизму. (Как и большинство занимающихся философией сознания американских философов, в общем и целом он остался глух к пост-куновской тематике философии науки, к идеям семантической несоизмеримости теорий, к социальной конструкции фактов, историцизму и др. Не потому, что эта тематика философски неинтересна, а потому, что его «инженерный» ум больше интересуют не метатеоретические вопросы, а решение конкретных философских проблем. Инженер применяет различные инструменты, но главным для него являются не инструменты, а понимание механизма и устранение его неполадок.) Он считает возможным обойти и инструментализм и релятивизм.

Дело в том, что понятие «интенциональная установка» у Деннета несет гораздо большую содержательную нагрузку, нежели интерпретация, прагматическая целесообразность и предсказательность. Будучи прагматической, подчеркивает он, в эпистемическом смысле она не является чисто инструментальной. Эта установка включает в себя также *нормативизм или норму рациональности*. Под нормативизмом обычно понимается установление стандартов или норм, согласно которым определяется (или конституируется) бытие объектов: что возможно для их бытия, а что невозможно. Например, для физической установки нормой является приписывание объектам законосообразного характера с учетом законов природы, для установки дизайна — функциональных и операциональных целей структуры объектов. *Нормой интенциональной установки является приписывание системам рационального поведения*. Это означает следующее. Во-первых, отношения,

215

которые интерпретатор приписывает объекту поведения, ограничены требованием, что объект в

целом следует представлять как имеющий при данных конкретных условиях верования, которые у него *должны быть*, и желания, удовлетворение которых было бы *в его интересах*. Во-вторых, рациональность здесь означает оптимальный дизайн, соотнесенный с целью или с оптимально взвешенной иерархией целей (у компьютера — это шах, выигранные фигуры, защита и т.д.) и рядом ограничений (у компьютера — правила и стартовые позиции, величина информации, имеющаяся в момент работы с шахматным полем). На основе холистического представления этих качеств вырабатывается наиболее разумное предположение относительно действий системы (в случае компьютера — возможных ходов). Когда речь идет о личности, это означает, что мы не можем приписывать интенциональные состояния отдельным частям личности, например мозгу или физическому поведению, а только всему комплексу «личность — диспозиции — цели — условия». В-третьих, интерпретатор приписывает объекту поведения *идеализированные* и эвристические нормы. То есть интенциональным системам, на практике подверженным ошибкам и сбоям, нормативизм приписывает нормы идеальной рациональности и истины.

Когда, опираясь на здравый смысл, мы объясняем и предсказываем поведение людей, говорит Деннет, мы исходим из посылки их рациональности. Обычно мы не ожидаем, что они будут реагировать иррационально на происходящие события, но когда они ведут себя таким образом, мы начинаем приспособлять к ним наши стратегии. Посылка рациональности столь глубоко укоренилась в наших привычках делать выводы (о людях), что когда наши предсказания оказываются ложными, мы сначала ищем средства для совершенствования условий получения информации (он, должно быть, не слышал сказанного, не знает английского языка, не видел то-то и то-то, не осознавал что-то и т.д.) или заново обдумываем цели, прежде чем ставим под вопрос рациональность системы в целом. В крайних случаях

216

поведение личностей может быть столь непредсказуемо с точки зрения интенциональной установки, что мы отбрасываем ее. В случае нерационального и непредсказуемого поведения личностей (например, психически нездоровых людей) мы отказываемся от интенциональной установки и обращаемся к установке дизайнера (ищем причины их поведения в неполадках организма).

То же самое относится и к предсказанию поведения животных. Преследуя зверя, охотник всегда приписывает ему интенциональность и хотя бы минимальную рациональность (или само-интерес по сохранению своей жизни), а затем корректирует объяснение поведения зверя, если он повел себя странно. Иными словами, приписывая атрибут интенциональности, следует настаивать на рациональности: встречаясь с явными нарушениями этого стандарта или нормы, сначала следует попытаться исправить нарушения (т.е. объяснить их в терминах, санкционированных самим этим стандартом), и если это не удастся, отбросить его. Согласившись сразиться с играющим в шахматы компьютером, можно зависить степень его рациональности, однако, по мере обнаружения изъянов в его интеллектуальных действиях, мы снижаем наши требования. В нормативистском понимании рациональности, говорит Деннет, состоит сущность всей науки и всего объективного знания.

Можно вообразить, что на какой-то другой планете живет отличный от нас народ. При допущении, что там действует дарвиновский отбор (а он обязательно должен действовать, коли происходит воспроизводство населения), существам этой планеты можно приписать верования и желания, следовательно интенциональность и рациональное поведение, хотя во всем остальном они могут совершенно не походить на землян. Возможен еще более «эксцентричный» вывод: мы вправе приписать этому народу *верование в логические истины*. Куайн был прав, когда говорил («Слово и объект»), что в любом поддающемся трансляции языке мы с необходимостью «откроем» наши логические связи. Подтверждениями наличия у

217

системы верований и желаний являются: потребность выживания, регулярность поведения и основанное на вере в естественный отбор допущение оптимального дизайна.

Возможен еще более «эксцентричный» вывод, а именно тот, что вера в логические истины

свойственна животным. Конечно, смешно представить мышь, размышляющую над примерами логических тавтологий, тем не менее эта картина не имеет отношения к делу, если мы исходим из «допущения, что нечто является интенциональной системой, есть допущение, что это нечто является рациональным. ... Поэтому безотносительно к тому, говорится или не говорится про животное, что оно *верит в истины логики*, следует предположить, что оно *следует истинам логики*. И действительно, наша мышь следует или верит в *modus ponens*, ибо мы приписываем ей следующие верования: а) *от меня налево находится кот и б) если кот находится налево от меня, мне лучше не идти туда*; а наше предсказание опирается на способность мыши делать выводы»<sup>1</sup>.

Правильно ли тогда говорить, что в любой интенциональной системе присутствуют все логические истины? Это не так. «Если бы система была идеальной или совершенно рациональной, — говорит Деннет, — в ней проявились бы все логические истины, однако любая актуальная интенциональная система будет несовершенной, и поэтому не все логические истины в виде верований следует приписывать любой системе. Более того, не все правила вывода действительной интенциональной системы могут быть правильными; не все ее выводимые верования могут быть истинами логики»<sup>2</sup>. Практика показывает, что каждой конкретной системе присущи сбои и ограниченности. Тем не менее этот факт ни в коей мере не подрывает принцип рациональности.

Тезис Деннета о том, что для понимания и объяснения поведенческих систем мы должны *приписывать* им интенци-

<sup>1</sup> Dennett D.C. *Brainstorms*. P. 10—11.

<sup>2</sup> *Ibid.* P. 11.

218

ональность и рациональность, вызвал бурю дебатов и острую критику с разных сторон. Прежде всего, со стороны интенциональных реалистов, видящих в интенциях внутреннее свойство сознания агентов. Их критика концентрировалась вокруг трех основных положений его теории. Во-первых, тезиса, что интенциональные состояния не являются «собственностью» агента интенциональности, не детерминированы присущими ему фактами и вообще не требуют сознания как такового. Во-вторых, предположения, что «приписывающее» понимание интенциональности, являясь произвольной конструкцией интерпретатора, тем не менее полагает рациональность и истину. В-третьих, что семантическое значение или содержательность сознания творится внешними для личности условиями. В книге «Интенциональность» (1983) Дж. Сёрль назвал сомнительной заявку Деннета на определение интенциональных состояний в терминах внешней установки: интенциональность следует считать проявлением ментальных состояний, по природе она присуща человеку внутренне, а не внешне. Именно сознание выступает первичным источником, наделяющим смыслами интенциональные установки и знаковую систему, с помощью которых достигается рациональность коммуникации: приписывание чему-то значения всегда соотносено с сознанием индивида — носителем языка — и производно от его внутренних намерений.

Критика последовала и со стороны «социологов», упрекающих Деннета за непоследовательность, связанную с попыткой натурализации интенциональности и подведения под интерпретивизм биологической основы (Р. Рорти, Бо Дальбом, Дж. Хогелэнд и др.). По мнению Рорти, нормативизм и интерпретивизм интенциональной установки заведомо предполагает отнесение ее к социальной конструкции. Он вообще предлагает Деннету отбросить тему «интенциональное и реальное». «Мой принципиальный совет состоит в том, чтобы ввести его заявление, что самость есть "центр нарративной гравитации", в контекст более общего требования, что все объекты сходны с самостями, будучи центрами *дескриптив-*

219

*ной гравитации*»<sup>1</sup>. Бо Дальбом считает серьезной теоретической ошибкой Деннета попытку перекинуть мосты между животными и людьми, биологическим и социальным с помощью интенциональной установки. Приписывая верования людям, мышам и компьютерам, он игнорирует социальное измерение и пытается биологизировать все феномены человеческого

познания и сознания. Ближе к истине, считает он, подошел Герберт Саймон, чья теория «искусственной науки» исходит из существования важных различий между биологией и инженерным искусством, а не ищет подобие между ними. Он советует Деннету быть осторожным и не описывать эволюцию, происходящую в результате естественного отбора, как процесс дизайна<sup>2</sup>. Деннет категорически отмечает и критику интерналистов, и критику экстерналистов-социологов. В ответ он пускает в ход свою тяжелую артиллерию — эволюционный аргумент и концепцию «свободно плавающей» рациональности (Free-Floating Rationality). Рассмотрим их.

Если эволюционный алгоритм работает, говорит он, практически полезные предсказания, основанные на приписывании системам интенциональности, должны иметь в природе достаточно прочные основания или резоны (reasons). Зададимся таким вопросом: каковы резоны функционирования природы, чьи они? Очевидно, что ничьи. Они никому не принадлежат, не проходят ни через чье сознание или «мудрость», они — «свободно плавающие» резоны, «дарвиновские резоны». Мы можем поставить вопрос по-другому: чьи резоны должны быть у организма, учитывая его эволюционную историю и конкретные возможности окружающей среды? Если допустить, что природа функционирует слепо, без «сознания» или «мудрости», по своим собственным алгоритмам,

<sup>1</sup> Rorty R. Holism, Intrinsicity, and the Ambition of Transcendence // Dennett and His Critics. Demystifying Mind. Ed by Bo Dahlbom. Cambridge Mass. 1993. P. 189.

<sup>2</sup> Bo Dahlbom. Mind is Artificial // Dennett and His Critics. Demystifying Mind.. P. 175.

220

мы должны допустить и «свободно плавающую» рациональность и интенциональность. Для функционирования интенциональных систем не требуется прохождения их через особый орган — сознание. Интенциональные содержания напрямую не детерминированы субъектами, чьи интенции рассматриваются. Интенциональность «плавает» столь свободно, что под давлением естественного отбора она может проявиться в системах, не имеющих сложной информационной организации. Ни играющий в шахматы компьютер, ни пчела не имеют никакого понятия об их собственных интенциональных состояниях, тем не менее они имеют их. Скалы, поскольку они не явились результатом естественного отбора для какой-либо цели, их не имеют. В системах со сложными информационно-процессуальными возможностями запрограммированные природой резоны могут осознаваться и на основе механизмов обратной связи модифицироваться. Человек, например, может преодолевать свои биологические цели, включая фундаментальные: некоторые люди, например, не хотят иметь детей.

Идея «свободно плавающей рациональности», считает Деннет, снимает второй упрек, а именно тот, что интерпретивизм не стыкуется с рационализмом. Действие естественного отбора, имеющего вектор рациональности, ограничивает релятивистский потенциал интерпретивизма. Интенциональным системам присуща «опробующая» способность, в процессе которой рождаются верования, и если в результате испытаний они оказались ложными, то отвергаются. Биологические виды «экспериментируют» с помощью мутаций, порождая огромное количество неэффективных систем, но ни одну из них нельзя называть системой-верованием именно в силу дефектов ее рациональности. Отсюда правомерен вывод, что *ложная «система-верование» есть концептуальная невозможность*. Подобно тому как растворимая в воде рыба есть эволюционная невозможность, система ложных верований, не поддающаяся когерентному описанию, — это концептуальная невозможность. У нас нет оснований приписывать какой-либо системе верования, если в целом они не соответствуют окружающим

221

условиям. Более того, способность к верованию не имела бы ценности выживания, не будь она способностью *верить в истины*. Такой же эволюционный (истинностный) критерий применим и к языковой коммуникации; способность к ложной коммуникации не будет вообще коммуникацией<sup>1</sup>. «Способность к коммуникации не укоренилась бы в эволюции, если бы она в общем и целом не была бы способностью к передаче истинных верований; последнее означает способность изменять других членов вида в направлении все более оптимального дизайна»<sup>2</sup>.



Таким образом, суть эволюционного аргумента Деннета в пользу рациональности и принципа истины сводится, во-первых, к тому, что, поскольку наши когнитивные способности эволюционируют, они должны быть подвержены отбору. С другой стороны, только когнитивные системы, в целом поддерживающие (удостоверяющие) истины, имеют шансы быть отобранными, поскольку ложные не будут иметь ценности выживания. Коли так, *принцип истины должен скреплять наши верования*. Для сферы интенционального конституитивной нормой является рациональное. «Рациональность, — повторяет Деннет, — есть мать интенции»<sup>3</sup>.

Дарвиновскими аргументами он парирует упреки, что значения у него оторваны от человека и переданы внешней среде. Проблема значения (meaning), как известно, является одним из «твердых орешков», с которым приходится иметь дело каждой аналитической теории. Понимание референциальности значения в конечном счете упирается в вопрос, зависит или не зависит приписывание какому-нибудь объекту значения от посылки существования внутренней интенциональности. Именно так считают Сёрль, Фодор и многие другие философы.

У Деннета свой путь к пониманию значения. Напомним, что традиционный путь к пониманию значения состоит в

<sup>1</sup> Dennett D.C. *Brainstorms*. P. 18.

<sup>2</sup> Ibidem.

<sup>3</sup> Ibid. P. 19.

222

рассмотрении отношения между именем и объектом, для которого оно используется в парадигмальном случае, а затем по аналогии понимается значение других слов. Радикальная альтернатива такому способу определения значения была предложена Витгенштейном (1953) в теории языковых игр. Согласно Витгенштейну, значения используемых слов или сигналов понимаются, в первую очередь, не в отношении с объектами, которые они обозначают, а из их функций внутри всего прагматического контекста использования слов.

Деннет в целом пошел по пути Витгенштейна в определении значения, но в отличие от него в контекст лингвистического поведения ввел фактор интенциональности. «Значение» находится у того, кто, исходя из интенциональной установки, приписывает его системе и предсказывает ее поведение; это аскриптивное понятие, функционально зависящее от всего контекста. «Значение» не есть конкретная часть систем, живых или неживых, а абстрактное теоретическое понятие, применяемое для целей предсказания точно так же, как понятие «масса» применяется к поведению физических объектов. Именно интерпретация формирует значение или суть того, что делает система, а также определенную способность предсказывать ее будущее поведение. Сторонники интенционального реализма, утверждающие, что значения, в отличие от артефактов, внутренне связаны с нашими ментальными состояниями, рисуют, убежден Деннет, мифологический образ первичной интенциональности; они полагают, что мы являемся особыми — не «означенными», а «означивающими» существами. У них значения с помощью интенциональности черпаются из скрытого внутреннего источника. «Доктрина первичной интенциональности утверждает, что, в то время как некоторые наши артефакты могут иметь производную от нас интенциональность, сами мы обладаем первичной (или внутренней) интенциональностью, произведенной не из чего. Аристотель говорил, что Бог — это Неподвижный Двигатель; данная доктрина провозглашает, что мы являемся «Никем не означенными означивателями» (Unmeant Meaners). Я никогда не верил в нее

223

и часто критиковал»<sup>1</sup>. На самом деле, настаивает Деннет, человека никто не наделял привилегией первичного источника значений: когда мы принимаем ментальные высказывания за осмысленные, мы просто принимаем по отношению к ним интенциональную установку. Короче говоря, на вопрос о «первичности» и «производности» следует смотреть с более широких дарвиновских позиций.

Дарвиновскими аргументами Деннет отвечает и на критику социологов, упрекающих его за натурализацию интенциональности и приписывание «верований» и людям, и мышам, и

компьютерам. Отвечая им, он поясняет, что такое «приписывание» не есть искусственный методологический прием. «Безотносительно к тому, каким бы термином мы ни называли их (верования — *Н. Ю.*), должны существовать информационные состояния, наличие которых позволяет нам предсказывать, манипулировать, объяснять с позиции интенциональной установки и летучую мышь и бармена, и именно это теоретически важно»<sup>2</sup>. Авторы, акцентирующие только социальный аспект интерпретивизма и нормативизма (Порти, Хогелэнд и Дальбом), обходят тему биологических возможностей их возникновения. «Но как иначе могли быть установлены социальные нормы... как не на плечах биологических норм...?»<sup>3</sup>. Только в существах, *уже* имеющих способность избегать ошибки или заниматься поиском истины (независимо от их знания или незнания этого), могут находиться основания для более самосознающей и софистичной когнитивной деятельности. «Над биологическим уровнем грубых (сырых) верований и сырых интенциональных икон человеческие существа возвели уровень социально сконструированных *объектов*, которые реплицируются, распределяются, вводятся в обиход, одобряются, ...отвергаются, ревизуются, атакуются, рекламируются, опровергаются. Докинс дал

<sup>1</sup> Dennett D. *The Intentional Stance*. P. 288.

<sup>2</sup> Dennett D. *Back from Drawing Board // Dennett and His Critics*. P. 229.

<sup>3</sup> *Ibidem*.

224

общее имя для них — меме...»<sup>1</sup>. Конечно, Дальбом, развивая идеи Саймона, прав, говоря о важных различиях инженерно созданных артефактов, с одной стороны, и артефактов, созданных процессом естественного отбора, с другой; между социологической и натуралистической перспективами есть напряженность. Тем не менее он не убедил, что эти различия перевешивают подобия; перспективы могут быть объединены с помощью иерархии установок<sup>2</sup>. Ключ к объединению Деннет видит в *отрицании* того, что показалось очевидным Герберту Саймону, полагавшему, что, вместо того чтобы исследовать отношения между артефактами и окружающими условиями, «наука об искусственном» должна концентрироваться на самих артефактах и их дизайне. На самом деле, биология, как и физика, является «естественной наукой» со своими «законами природы» и «естественными видами». Сводить ее к «искусственному» вряд ли возможно.

По ложному пути, считает Деннет, пошли не только интенциональные реалисты, социологи, но и элиминативисты, выбросившие интенциональность из объяснительных схем и признавшие ее негодным понятием для объяснения рациональности. Скиннер посчитал теории, преподносимые в интенциональных идиомах, эмпирически пустыми и поэтому не могущими служить фундаментом для психологии. Он прав только частично. Интенциональная теория не является единственным фундаментом психологии, поскольку она *предполагает, а не объясняет* рациональность или интеллектуальность<sup>3</sup>. Однако она не покажется эмпирически пустой, если обратить внимание на выполнение ею *предсказательных* функций, которые невозможно реализовать на основе жесткого скиннеровского бихевиоризма. Известен факт, когда тысяче математиков предложили умно-

<sup>1</sup> *Ibid*. P. 230.

<sup>2</sup> Эту тему Деннет развивает в статьях: «Artificial life: A Feast for Imagination» (1990), «Cognitive Science as Reverse Engineering: Several Meanings of "Top-down", "Bottom-up"».

<sup>3</sup> Dennett D. *Brainstorms*. P. 13.

225

жить 5 на 7, из них более 900 ответили правильно, а остальные по ряду причин (не расслышали, не поняли задание и др.) дали неверный ответ. Данный факт объясним с точки зрения «психологии» интенциональных систем, но не с точки зрения старой механистической и эмпирической психологии, ограничивающей себя исследованием «входа» и «выхода» информации. Скиннер попытался обойтись без интенционального языка, предсказывая телесные реакции на чисто физические стимулы, однако это не дало ему возможности сформулировать альтернативную ментализму позицию относительно объяснений, предсказаний и рациональности. Являясь пустой как психология — в силу *посылки* рациональности, *а не строгого логического или*

*эмпирического ее доказательства*, — интенциональная теория, говорит Деннет, не является пустой с точки зрения ее прагматических и предсказательных возможностей. Например, теория игр, будучи не научной психологией, а формальной нормативной теорией, по необходимости является интенциональной. Применяемые ею к человеческим субъектам теоретико-игровые предсказания достигают точности в силу эволюционной гарантии, потому что дизайн человека-игрока спроектирован как особый случай рациональности. То же самое относится и к социальной науке экономике, обладающей огромной предсказательной силой. Не являясь психологической теорией, она предполагает то, что психологии еще надлежит объяснить. Экономические объяснения и предсказания в ней являются интенциональными и осуществляются масштабно потому, что в общем и целом отдельные люди выступают хорошими приближениями к оптимальному оператору в сфере рынка<sup>1</sup>. «Политическая наука, социология, антропология и социальная психология... применяют интенциональный, "ментальный" или "когнитивный" словари — говорят о "верованиях", желаниях, ожиданиях, распознавании, действии и т.п.», и это дает хороший эффект<sup>2</sup>.

<sup>1</sup> Dennett D. *Brainstorms*. P. 16.

<sup>2</sup> *Ibid.* P. 15—16.

226

#### 4. Уровни интенциональности и интеллектуальности

Обратимся теперь к аспекту теории интенциональности Деннета, связанному с его объяснением происхождения человеческого типа интеллектуальности. В предыдущих главах мы уже говорили о том, что он резко противопоставил лингвистически вооруженное сознание человека психике животных, тем самым приняв вектор социологизма. На самом деле у него более сложный ход рассуждения. При решении этой проблемы, как и при решении всех других проблем, он прибегает к двойной перспективе и кружным путем погружает человека в глубины природы.

При объяснении информационно-когнитивных способностей природных систем важно учитывать возможность приписывания им разных *уровней интенциональности*. Можно приписывать им интенции нулевого, первого, второго и третьего порядка. Деннет образно именуется их «дарвиновскими», «скиннеровскими», «поппернианскими» и «грегорианскими» системами (последние названы по имени психолога Ричарда Грегори, исследовавшего влияние артефактов на становление интеллектуальности). Базисными являются простые «дарвиновские» системы (растения, низшие животные), чьи экологические приспособительные реакции в стремлении к выживанию фиксированы эволюцией. Когда в обычной речи мы объясняем поведение дерева с помощью фраз «оно стремится к свету», «оно испытывает жажду», мы применяем по отношению к нему интенциональную установку, позволяющую в узких пределах осмысливать, контролировать и предсказывать его поведение без знания каузальных контролируемых механизмов. Даже если эта установка обладает малой эвристической ценностью, более мудро принять дерево за неразвитую, малоинтересную, незначительную интенциональную систему, нежели обозначить предел, только выше которого правомерны интенциональные интерпретации. Следующий уровень

227

«скиннеровские» системы. Они способны усваивать новые стратегии и обучаться новым простым реакциям путем подкрепления (через поощрение) поведения. «Поппернианские» существа уже способны задействовать дополнительный и более мощный ресурс: они используют некую внутреннюю модель их мира, дающую возможность предварительно опробовать в воображении поведенческие акты до их физической реализации. Эта стратегия используется большинством животных, чья софистичность превосходит когнитивные способности простых беспозвоночных животных. Учеными высказаны разные мнения относительно интеллектуальности высших животных. Некоторые считают, что, поскольку у собак, шимпанзе и других высших животных нет языка, они не могут подняться до интенциональных систем второго порядка. Такое мнение Деннет считает ошибочным. Сложное поведение животных можно отнести к более высокой

интенциональной системе — системе второго порядка, то есть к системе, которой мы приписываем не только простые верования, желания и другие интенции, но и верования и другие интенции *относительно* верований, желаний и других интенций. Интенциональная система «S» будет интенциональной системой второго порядка, если среди приписываемых ей свойств есть такие как: «*S* верит, что *T* желает, что *p*»<sup>1</sup>. (См. в «Условиях личностного» примеры с собакой Фидо, выпрашивающей еду, с собакой Эшли, «обманывающей» хозяина, и птицей, имитирующей сломанное крыло, чтобы увести от гнезда преследователя.) Наличие в эволюции птиц такого тропизма, как *обман* преследователей или внушение ложных верований, трудно объяснить без приписывания им интенций второго порядка. Однако это не дает основания считать, что у шимпанзе, собак, дельфинов и других животных есть признаки рудиментарной сознательности. Стремление птицы внушить преследователю ложное верование не имеет ни пространства, ни

<sup>1</sup> Деннет Д. Условия личностного. С. 457. (Приложение 1 настоящего издания.)

228

необходимости для репрезентации данной софистичной интенции в «мыслях» или «сознании», ибо «ни птица, ни эволюционная история, ни Мать-Природа не имеют нужды мыслить эти мысли для подкрепления нашего объяснения»<sup>1</sup>. И вообще, мышление о мыслях не делает подлинно интеллектуальное поведение интеллектуальным.

И, наконец, интенциональные системы третьего порядка, которыми обладают «грегорианские» существа — люди. В них усматриваются все присущие низшим уровням свойства, плюс еще «добавка» — язык и культура, практически безгранично расширившая их когнитивное пространство. Язык и культура играют двунаправленную роль в возникновении феномена интеллектуальности человека: с одной стороны, они требуют для своей работы интеллектуальность, с другой стороны, интеллектуальность для своей реализации требует задействованности этих инструментов.

Принятие этого общего тезиса еще не проясняет специфику человеческой интеллектуальности. В традиционной философии было принято считать определяющими характеристиками человека разумность и самосознание. Современные философы склонны видеть специфику человека в языковой коммуникации. Деннет вносит в эти представления существенную коррективу. И самосознание, и лингвистическая коммуникация являются необходимыми, но недостаточными условиями. Чем бы мы ни считали личность — душой, самосознающим моральным агентом, эмерджентной формой интеллектуальности, вместительным языковым коммуникаций, — ее, в первую очередь, следует рассматривать как интенциональную систему, и это детерминирует все ее остальные проявления. Человек — это «подкласс интенциональных систем, имеющих язык и способных вступать в коммуникацию; только это дает рамки для теории сознания»<sup>2</sup>. Сначала следует разобраться с интенциональностью, а затем со всем

<sup>1</sup> Там же. С. 463-464.

<sup>2</sup> Dennett D.C. Brainstorms. P. 17.

229

остальным — сознанием, языком, значениями, моралью и т.д. (Поэтому у Деннета теория интенциональности предшествует теории сознания.)

Конечно, без наличия у интенциональной системы способности к лингвистической коммуникации невозможно приписывать ей верования, рациональность и т.п. Язык является главным условием интерпретативной деятельности. При этом не следует забывать, что сам язык является продуктом эволюционного процесса. «Способность к языку, без всякого сомнения, является венчающим эволюцию достижением, достижением, которое, в свою очередь, питает ее для производства множества других сложных и рациональных систем, тем не менее все они могут рассматриваться как формы адаптации, которые сами являются объектами тех же самых условий полезности для окружающей среды, как и всякий другой поведенческий талант»<sup>1</sup>.

Когнитивным системам животных тоже свойственны и коммуникативность, и то, что называется взаимностью или взаимопониманием. Однако для интенциональности второго уровня они жестко фиксированы эволюцией. Только интенциональность систем третьего уровня преодолевает

тропизм эволюции. К определению интенций *третьего порядка* ближе всего, считает Деннет, подошел Х. Грайс в его попытке определить то, что он называет «ненатуральным» значением (1957). Это наблюдается, когда произносящий придает значение чему-то путем высказывания о нем в терминах интенций произносящего. Первоначальная дефиниция Грайса выглядит следующим образом.

«*U* (например, преподаватель) имеет в виду что-то, произнося *X*», и это является истинным, если обращаясь к *A* (аудитории слушателей) *U* произнес *X* с намерением, чтобы:

1. *A* отреагировала на произносимое определенным откликом *R*;

2. *A* думала (признала), что *U* имеет намерение (1);

3. *A* выполнила (1) на основании выполнения (2).

<sup>1</sup> Dennett D.C. *Brainstorms*. P. 17.

230

У Грайса интенция (2) приписывает *U* интенцию не только второго, но и третьего порядка. Преподаватель, например, должен иметь намерение, что его аудитория распознает его намерение, чтобы у нее возникла определенная реакция на то, что он говорит. К интенциям третьего порядка относятся все случаи обмана и манипуляции говорящего своими слушателями, основанные на эксплуатации рациональности жертвы.

Деннет подчеркивает, что в ситуации обмана или манипуляции, находясь в сложных коммуникациях, обманщик и его жертва не всегда проводят свои рассуждения через осознанный мыслительный процесс. Человек, вступающий с другими людьми в различного рода коммуникации, не осознает, что он проделывает такие сложные мыслительные ходы. Если в качестве примера взять систему человек—компьютер, коммуникация человека с компьютером не зависит от способности машины «в мыслях» распознавать хитроумные интенции человека. Это вовсе не означает, что в будущем все коммуникации человека с компьютерами сохранят эту интеллектуальную ограниченность. Это только значит, что сегодня мы не вступаем в интеллектуальную коммуникацию с компьютером в строгом (грайсовом) смысле слова.

## 5. Первичная и производная интенциональность. Мать-Природа как первичный дизайнер

Интенциональная установка — это только установка, то есть часть объяснительной схемы. Оригинальная и интересная часть работы Деннета состоит в проекции картины мира, лежащей за этой схемой и оправдывающей ее как объяснительную схему. Речь идет о его концепции первичной интенциональности Матери-Природы и роли естественного отбора как движущего фактора ее жизнедеятельности.

231

Как всякий ориентированный на науку человек, Деннет считает сознание, интенциональность, способность к языку и производству значений результатами эволюционного развития природы. Но в отличие от многих других философов натуралистического склада он делает более решительный шаг: предлагает рассматривать саму природную эволюцию с точки зрения интенциональной установки, а еще точнее — видеть в ней имманентное развертывание интенциональной установки. Наша интенциональность, говорит он, в конечном счете, коренится в «Матери-Природе» или процессе эволюции, именно она — Главный Дизайнер.

Для пояснения этого тезиса Деннет вводит понятия *первичной* (или исходной) интенциональности (Original Intentionality) и *производной* интенциональности (Derived Intentionality). Первичная интенциональность, или «прото-интенциональность», свойственна только природе. Интенциональность человеческих существ вторична: это своего рода *артефакт Природы*<sup>1</sup>. Творя значения, артефакты и культуру, человек, сознательно или бессознательно, исходит из уже зафиксированной эволюцией — через естественный отбор — прото-интенциональности или прото-значений. В свою очередь, по отношению к языку и артефактам его интенциональность первична: дорожные указатели, карты, слова, теории и другие знаковые системы несут на себе



печать его интенциональных установок и интерпретаций. Словом «карта» обозначают местность, дорожными знаками — правила движения и т.д. Эти маркировочные знаки вторичны по отношению к нашим намерениям. Тем не менее наша интенциональность не является исходно первичной, наделяющей значениями карты, знаки и т.д. В конечном счете она производная от жизнедеятельности и алгоритмов Матери-Природы<sup>2</sup>.

Более конкретными дизайнерами интенциональности являются бессловесные гены, которые каждый несет в себе. «Поэтому наша интенциональность выводима из интенционально-

<sup>1</sup> Dennett D. *The Intentional Stance*. P. 300.

<sup>2</sup> Ibid. P. 288.

232

сти «эгоистичных» генов. Именно они являются Unmeant Meaners, а не мы сами<sup>1</sup>. Конечно, в аналогии с генами есть и дизаналогия. «Немного необычно представлять гены в виде умных дизайнеров; тем не менее сами они не могут быть более глупыми, они просто являются благодеяниями процесса дизайна. Тогда кто, или что, осуществляет дизайн? Мать-Природа, конечно, точнее — длительный, медленный процесс естественного отбора<sup>2</sup>. Работа генов аналогична работе инженера-проектировщика, они выступают опосредующими детерминаторами в генезисе интенциональности, носителями чего-то, конечный источник которого лежит вне их самих — в природе. Отсюда следует вывод, что если Мать-Природа является «дизайнером» наших генов, нам следует принимать ее за очень большую интенциональную систему. Для подтверждения этого вывода требуются только две вещи: во-первых, чтобы сохранялась напряженность динамики эволюции и, во-вторых, чтобы в ней продолжал действовать вектор рациональности.

Разумеется, эта, по сути метафизическая, идея Деннета не была встречена с пониманием ни учеными-биологами, ни популяризаторами науки. Эмпирически мыслящие ученые предпочитают объяснять эволюцию взаимодействием чисто каузальных факторов. Они критикуют его за придание естественному отбору первостепенной значимости и за игнорирование фактов производства эволюцией нежизнеспособных видов. С нашей точки зрения, они спорят с ним как с ученым-эмпириком, в то время как он — философ и его концепция является философской. А теория эволюции, учитывая наличие в ней огромного количества белых пятен, не может обойтись без философских гипотез.

Полемизуя с учеными (Стивенем Гоулдом и др.), Деннет оговаривает, что рациональность Матери-Природы, как и рациональность человека, не следует понимать в смысле отсутствия ошибок и накладок. «Когда естественный отбор отбира-

<sup>1</sup> Ibid. P. 298.

<sup>2</sup> Ibid. P. 299.

233

ет дизайн, он может "выбрать" специфический дизайн по одной, нежели другой причине<sup>1</sup>. Игра каузальных факторов может приводить к эволюционным сбоям, мутациям, случайностям, появлению нежизнеспособных видов. Подобным же образом сбои и ошибки происходят и в развитии человеческого знания, и в жизни отдельного индивида, когда те или иные события ломают его планы и перспективы на будущее. Рациональность природы — это идеализированная рациональность. Точно так же, как мы идеализируем рациональность когнитивных систем человека, принимая по отношению к нему интенциональную установку, мы идеализируем рациональность эволюционного дизайна природы, принимая ее за интенциональную систему. Никто не отрицает, что многие события в жизни Земли были случайностями, результатами цепи физических причин (гибель динозавров и образование ниши для последующего развития млекопитающих, возможно, были связаны со случайным падением астероида). Эти причины объяснимы с позиции физической установки. Объяснить физическими причинами вектор эволюции не удастся. Нельзя считать, что объяснение на основе игры физических каузальных факторов по своему влиянию способно конкурировать с объяснением на основе естественного отбора. Это можно подтвердить чисто статистически. Самое главное, что упускают из виду

критики, — это разные уровни объяснительной организации. Каузальное объяснение делается на уровне физической установки, целеориентированное — на уровне интенциональной установки. (Этой теме посвящен раздел «Прочитывая сознание Матери-Природы» в книге «Интенциональная установка»<sup>2</sup>.)

В связи с философским подходом Деннета к эволюции нужно заметить следующее. И исторически и онтогенетически процесс естественного отбора у него выступает основой *целенаправленности*: последняя делает возможной осмыслен-

<sup>1</sup> Dennett D. *The Intentional Stance*. P. 299.

<sup>2</sup> Ibid. P. 298—300.

234

ность и содержательность интенциональных систем. Поскольку естественный отбор понимается как отбор способностей, служащий каким-то целям, это значит, что в картину мира вносится *телеологизм*. Без большой натяжки можно сказать, что естественная телеология выступает у Деннета *базисным логическим стандартом*, на фоне которого эволюционирует осмысленность, проявляясь в содержании мыслей и целенаправленном поведении личностей.

В предыдущих главах мы уже говорили о том, что Деннет воспринял ряд идей социобиологов о коэволюции биологического и социального. Суть его эволюционизма тем не менее сложнее, нежели просто энтузиазм в отношении идей коэволюции. Будучи не только дарвинистом, но и тьюрингианцем, он утверждает, что *сам естественный отбор алгоритмичен*, поэтому существует исчислимость, более того, как бы «вычислимость» (хотя он и использует это слово в кавычках и в сопровождении «как бы») на всем пути в прошлое — от нас к нашим ранним молекулярным предшественникам. То есть ему видится алгоритмичность (и телеология) не только в работе «Машины Дарвина», но и в искусственном интеллекте компьютера, и в мыслительных шагах человека. Такое видение свидетельствует о том, что его собственная «максимальная метафизика» существенно отличается и от метафизики социобиологов, и от метафизики тьюрингианцев.

## 6. Джерри Фодор о телеологизме эволюционного аргумента Деннета

Как мы уже сказали, концепция интенциональности Деннета является предметом острых дебатов. Специалисты многих дисциплин высоко оценивают ее методологическую прикладную значимость. В то же время ее философский смысл подвергается атакам. Прежде всего, со стороны тех, кто считает интенциональность проявлением особого внутреннего

235

свойства сознания, а не установкой интерпретивизма. Деннетовскую позицию они относят к инструментализму, говорят о том, что его стремление подменить первичную интенциональность человека производной интенциональностью знаковых систем есть еще одна попытка изъять сознание из объяснительных схем. А его обращение к глобальному «дизайнеру» интенциональности — Матери-Природе — следует расценивать как спекулятивный маневр, который невозможно ни доказать, ни опровергнуть.

Деннета, безусловно, глубоко задевают эти атаки. Он продолжает настаивать, что «приверженность доктрине первичной интенциональности есть исходный источник путаницы в современной англо-американской философии сознания... на самом деле... интенциональность наших ментальных состояний и процессов столь же производна, как интенциональность книг и карт (и внутренних состояний наших роботов)»<sup>1</sup>. Конечно, говорит он, идею первичной «внутренней» интенциональности защищают такие авторитетные философы, как Д. Фодор, Дж. Сёрль, Ф. Дречке, С. Крипке, Д. Чэлмерс и другие, однако у них есть не менее выдающиеся оппоненты — У. Куайн, У. Селларс, П. Черчленд, Р. Рорти, Столкнер, Д. Дэвидсон, Дж. Хогелэнд, Р. Милликэн, М. Минский, Д. Хофстадтер и почти все занимающиеся исследованием Искусственного Интеллекта. И не факт, что первая группа ближе стоит к истине.

В 1993 году был опубликован сборник статей «Деннет и его критика. Демистифицируемое сознание», выпущенный к его пятидесятилетию (он выполнен в виде анализа его взглядов и ответов на критику)<sup>2</sup>. Один из важнейших пунктов несогласия участвующих в сборнике

философов и ученых касается толкования *отношений физики, биологии, психологии,*

<sup>1</sup> Dennett D.C. and Haugeland J. Intentionality. Internet: <http://ase.tufts.edu/cogstad/pubpage.htm>. P. 8.

<sup>2</sup> Dennett and His Critics. Demistifying Mind. Ed. by Dahlbom Bo. Oxford, Cambridge, 1993. P. 76. Далее: Dennett and His Critics. P. ...

236

точнее, статуса применяемых в этих науках объяснений. Позиция Деннета такова: психология и биология существенно отличаются от физики; в то время как в биологии и психологии объяснения с позиции интенциональной установки являются нормативными, основанными на допущении рационального поведения живых существ, объяснения в физике не являются нормативными и опираются на законы природы.

Джерри Фодор и Эрнст Леплор не согласились с такой позицией. С их точки зрения, психологии в определенной мере можно приписать ранг науки; многие применяемые в ней интенциональные объяснения аналогичны объяснениям других наук. Например, объяснение иллюзорного восприятия Луны в физике вполне укладывается в интенциональное объяснение. Но главный пункт их несогласия с Деннетом — приписывание эволюционному аргументу основания принципов рациональности и истины. Они находят сомнительным тезис, что, поскольку наши когнитивные способности эволюционируют, они должны быть подвержены действию естественного отбора, при этом отбираются только те когнитивные системы, которые в целом поддерживают (удостоверяют) истину. Система, являющаяся результатом отбора, говорят они, не требует, чтобы все ее подсистемы имели ценность выживания; некоторые из них могут быть рудиментарными и не иметь ценности выживания. И вовсе не очевидно, что диспозиция верить преимущественно в истинное обладает соревновательным преимуществом в сравнении со способностью верить преимущественно в ошибочное. Если бы это был реальный факт, эпистемологи раз и навсегда получили бы возможность покончить со скептицизмом в отношении познания<sup>1</sup>.

Существенное логическое противоречие они усматривают в том, что нормативизм у Деннета включает в себя интерпретивизм и в то же время предполагает принцип истины и принцип честности. Интерпретивизм, говорят они, означает, что у агентов реально имеются не верования и желания, а

<sup>1</sup> Ibid. P. 74.

237

только установки, приписывающие им верования и желания. Поэтому Деннету, утверждающему, что верования и желания реально не присущи агентам, не поможет апелляция к эволюционизму и к борьбе за существование для сохранения истины. Именно по этой причине большинство философов, придерживающихся эволюционной стратегии в отношении интенциональности, в отличие от Деннета являются реалистами в отношении содержания сознания.

Деннету мало помогает, считают Фодор и Леплор, апелляция к Матери-Природе. Они предлагают рассмотреть следующее мысленное допущение. Предположим, что экологическая функция лесных пространств состоит в предотвращении эрозии верхнего слоя почвы; это та функция, для чего леса *предназначены* покрывать почву, или в этом их цель. Однако в таком случае экология либо подтверждает понятие функции, либо не подтверждает. Если подтверждает, тогда функция леса покрывать почву есть дело факта; если не подтверждает, тогда мы не сможем улучшить объяснение, приняв «интенциональную установку» по отношению к эрозии. Мы, конечно, можем придумать сказку о том, что Отец-Эрозия стремится смыть верхний слой почвы, а Волшебница Леса хочет воспрепятствовать этому. Сказка может быть полезной для мнемонических целей и развлечения детей, однако такого рода история не может быть использована в качестве доказательства существования или несуществования экологических функций<sup>1</sup>. Свой мысленный эксперимент они заканчивают вопросом: если история об Отце-Эрозии не делает законным интерпретивизм в отношении функций в экологии, тогда почему история о Матери-Природе должна узаконить интерпретивизм в отношении функций в биологии? Общий их вывод таков:

Деннет являет собой пример того, «как боги наказывают инструментализм: отказавшись отличать теории от сказок, мы вскоре теряем способность отличить сказки от теорий». Мораль из

<sup>1</sup> Dennett and His Critics. P. 76.

238

сказанного следующая: «Если приписывание интенционального должно пониматься в терминах эволюционного объяснения, тогда вопрос о сохранении принципа истины является скорее эмпирическим, нежели концептуальным»<sup>1</sup>.

Деннет, разумеется, не согласился с такой трактовкой его позиции и вообще с их взглядами на характер философской работы. У Дж. Фодора и Э. Леплора, говорит он, «философия искореженная и четвертованная»<sup>2</sup>. В общей форме его ответ можно представить так: авторы пошли по традиционному пути и применили к его идеям *логические требования*: они сначала представили их в виде логических аргументов, а затем бьют с помощью истории об Отце-Эрозии и Волшебнице Леса. На самом деле, разъясняет он, рассуждения об основаниях (reasons) Матери-Природы вовсе не нацелены служить логическими аргументами, чтобы узаконить интерпретивизм; их цель другая — *объяснить* интерпретивизм, сделать его постигаемым и приемлемым. Именно так это принято в науках, в биологии или физике, когда у объяснений не всегда имеются логические или эмпирические основания (reasons). В этих случаях можно говорить об «основаниях», но они будут «свободно плавающими основаниями» (free-floating rationales), *ничьими основаниями*, основаниями, которые сами входят *в алгоритмы природы*. Одним словом, единственно открытый путь к приписыванию атрибутов в науке — это санкционировать эти «свободно плавающие основания»<sup>3</sup>. Если же оставаться на уровне чисто логических аргументов, как это делают Фодор и Леплор, навести мосты через обозначенную Brentano пропасть никогда не удастся.

После выхода в свет книги Деннета «Опасная идея Дарвина» Фодор откликнулся на нее рецензией, в которой отнес его философию к «амбициозному» варианту метафизики, построенной на телеологическом понимании адапционизма и на-

<sup>1</sup> Ibid. P. 77.

<sup>2</sup> Ibid. P. 215—217.

<sup>3</sup> Ibid. P. 215.

239

турализации интенциональности<sup>1</sup>. У Деннета, пишет он, нет ясности относительно различия *методологического* и *онтологического* адапционизма. Первый обычно относят к приемам, принятым в поведенческих науках. Второй смысл есть тезис, что формирование организмов происходило главным образом благодаря естественному отбору. Этот тезис неверный: многое в эволюционном развитии является результатом различных форм канализации в морфологии, нежели соревновательности среди разных дизайнов, которые сформировались псевдослучайно. Попытка Деннета использовать установку дизайнера в отношении естественного отбора как основы для натурализации интенциональности в целом основывается на телеологизме. Понятие «отбор для» у него совмещается с понятием «отбор для» в соответствии с «чем-то вроде понятия в сознании». Эта попытка безнадежна, поскольку бессознательность и близорукость Матери-Природы слишком велики. Она никогда не предпочитает одно другому, иначе как на основе прямых или косвенных каузальных взаимоотношений. И в этом ее существенное отличие от людей, полагающих свои цели осознанно. Вывод Фодора таков: не существует никакой естественной телеологии, и вообще «не существует никакой Матери-Природы»<sup>2</sup>.

Деннет считает, что Фодор не понял сути его позиции и смешивает эмпирический и философский уровни понимания. Естественный отбор есть организационный принцип, *концептуально эквивалентный* принятию установки дизайнера по от-

<sup>1</sup> Он пишет: «Деннет предлагает вариант метафизической конструкции, в которой (нео)дарвинистское объяснение адаптации кладется в основу теории естественной телеологии; и данная теория естественной телеологии, в свою очередь, кладется в основу объяснения значения символов и интенциональности сознания. Эта программа, конечно, в высшей степени амбициозна, даже значительно больше, чем адапционизм per se». Fodor J. Deconstructing

Dennett's Darwin. In: Fodor J. Critical Condition. Cambridge MA, 1998. P. 171.

<sup>2</sup> Ibid. P. 186.

240

ношению к эволюции. Отношение между актуальными понятиями и примерами селекции в истории жизни аналогичны отношению между событиями мозга и интенциональными состояниями, то есть отношениям между *illata and abstracta*. Иначе говоря, естественный отбор — это понимание перспективы эволюции с помощью абстрактных паттернов. Фодор ошибается, думая, что интенции являются микрокаузальными силами на уровне *illata*. Суть состоит в том, что интенции могут существовать как реальные паттерны, без необходимости быть подобными каузальным столкновениям бильярдных шаров. Интенции не оказывают каузального воздействия ниже их собственного уровня *abstracta*, и поэтому, конечно, этого не делает и Мать-Природа. Разумеется, ее планы и цели видимы только ретроспективно, но то же самое относится и к нашим планам, поскольку они проявляют себя только в нарративах, которые рассказывают о себе люди, и когда по отношению к ним принимается интенциональная установка.

\* \* \*

Заключая главу, выскажем ряд соображений по поводу деннетовской концепции интенциональности. Из представленного материала читатель может судить, что Деннет возложил очень большую миропонимающую и связующую нагрузку на понятие «интенциональность». По его собственному признанию, «это слишком многое, что можно ожидать от одного понятия, но несколько не меньше, чем сам Brentano ожидал в период, когда во времена менее фрагментированной науки ввел интенциональность в виде признака, самым фундаментальным образом расколовшего Вселенную: размежевавшего ментальное и физическое»<sup>1</sup>. По его замыслу, новое содержание этого понятия (интенциональные системы, интенциональная установка, уровни интенциональности, первичная и производная интенциональность и др.) является радикальной альтернативой тезису Brentano. Оно позволяет нарисовать

<sup>1</sup> Dennett D. Brainstorms. P. 22.

241

цельную организмическую картину бытия без провалов, пропастей и демаркационных линий, представить мир в виде иерархии интенциональных уровней, в соответствии с которыми понимаются взаимодействия неживой природы, живой природы, человека и артефактов. С помощью этого понятия, убежден он, можно перекинуть мосты, связывающие физические науки со сферой интенционального (куда он включает представления здравого смысла, научную рациональность, теорию игр, объекты биологии, психологии, этики и др.). С его же помощью можно ликвидировать пропасть между естественным интеллектом и искусственным. Предполагается, что данная теория избегает ловушки редуционизма: разные интенциональные системы не редуцируются либо к физическому, либо к компутационному уровню, но в то же время вполне с ним совмещаются.

Удалось ли Деннету реализовать свою задумку? Однозначно ответить на этот вопрос трудно. Нам представляется, что при всех зафиксированных критиками изъянах деннетовской теории интенциональности ее общая стратегия, а это социобиологическая стратегия, нацеленная на фиксацию «пунктов встречи» человеческого и природного, является новаторской и перспективной (если сравнивать ее, например, с логицизмом Фодора, социологицизмом Рорти или биологицизмом Черчленда). Заслуживает всяческого внимания методология эволюционного когнитивизма, выдвинутая как альтернатива традиционному эмпиризму и логицизму в понимании рациональности. Конечно, эта методология, по-другому эволюционная эпистемология, имеет уже давнюю традицию. В ее рамках работали Ж. Пиаже, К. Поппер, И. Лакатос, С. Тулмин, Д. Кемпбел и другие. Тем не менее Деннет внес в эту традицию свой вклад, связанный, прежде всего, с теорией интенциональности. Разумеется, что «работа на стыках» и эволюционный когнитивизм уязвимы для логической и эмпирической критики. И все же он не отступает от своих принципов. Решение извечных дилемм философии ему видится в отказе от дедуктивного постро-

242



ения теорий, от принципа «или-или», в принятии позиции концептуального компромисса. Предлагаемую им эволюционистскую и прагматическую интерпретацию рациональности мы бы расценили как вариант «мягкой рациональности», если сравнивать этот вариант с «жесткими» логицистскими типами рациональности (неопозитивизмом, теорией Дж. Фодора и др.). Он прав, если иметь в виду итоги длительного спора об основаниях рациональности, в том отношении, что поскольку ни одна интенциональная система не является идеально рациональной, то требование оптимальной рациональности нереально. Жесткие критерии здесь не работают. На каких-то этапах имеет смысл прекратить проверять допущение, что та или иная система является рациональной, и согласиться, что система определена достаточно полно, для того чтобы приписывать ей разумные верования.

Конечно, критики правы, что деннетовская концепция интенциональности и рациональности привязана к определенной интерпретации эволюционизма, которую для краткости называют «адаптационизмом», то есть принимающей за главные движущие факторы развития естественный отбор и адаптацию к среде. Если исходить из иной трактовки (а их немало), тогда его концепция теряет силу. Однако природа эволюции — открытый вопрос, и философские гипотезы здесь более чем уместны. (Вспомним, что К. Поппер отнес дарвинизм к «метафизической исследовательской программе», а его концепция фальсификационизма, построенная на идее целенаправленности естественного отбора и адаптационизме, была предметом атак со стороны и ученых, и философов.) Конечно, замечание Фодора, что концепция Деннета вносит в мир телеологию, имеет под собой почву. «Установка дизайнера» Деннета — это телеологическая установка, а придание Матери-Природе статуса «Главного Дизайнера» подтверждает вывод о его приверженности естественной телеологии. По этому поводу можно только заметить, что без идеи телеологизма проблематично ответить на вопрос, почему естественный порядок природы

243

таков, что создалась возможность для появления рационально мыслящего и целеполагающего человека<sup>1</sup>.

Теорию интенциональных систем Деннета в определенной мере можно расценить как введение новой — «нейтральной» — сферы без традиционной «ментальной» нагрузки. В связи с этим возникают вопросы относительно онтологии нейтрализма, путь к разрешению которых неясен. Прежде всего потому, что в его объяснительной схеме остается центральным различие между физическим бытием и бытием, обладающим интенциональностью. При всем том, что физическая установка, установка дизайнера и интенциональная установка уравниваются в своих правах — являются прагматическими стратегиями, — в онтологическом смысле они не одинаковы. Две первые установки наделены онтологическим статусом, у интенциональной ее нет. Физические сущности являются *illata*, то есть выводимыми сущностями, а интенциональные — *abstracta*, то есть постулируемыми сущностями. Сфера интенционального хотя и совместима с физикой, но не редуцирует-

<sup>1</sup> «Аргумент от дизайнера», или телеологический аргумент, имеет давнюю историю. Он использовался Фомой Аквинским для доказательства бытия бога: имеется в виду его апелляция к упорядоченности, целесообразности, интеграции природы. После Дарвина, введшего в обиход понятие «естественный отбор», это понятие составило один из главных принципов естественной телеологии, в которой, в отличие от религиозных учений, и физические причинно-следственные отношения, и целенаправленность помещены в саму природу. В связи с этим можно заметить, что Мать-Природа выполняет у Деннета примерно такую же роль, какую традиционное христианство отводило Богу. У Декарта истинность и рациональность наших верований удостоверяет благожелательный Бог. Существа, которые не следуют божественно установленному естественному порядку вещей (или имеют адаптационные изъяны, как говорят дарвинисты), обречены на исчезновение. Главное, что и в христианстве, и у Деннета естественный порядок мыслится как целенаправленный, то есть телеологически.

244

ся к ней. В проведении этих различий, как нам кажется, таится угроза возвращения к «тезису Brentano»<sup>1</sup>.

Последнее наше замечание является метафилософским. Мы уже говорили, что, восприняв у

Куайна натуралистический подход к эпистемологическим проблемам, Деннет воспринял и куайновские тезисы о неопределенности значения, аргументы против дихотомии теоретического и эмпирического, аналитического и синтетического и др. (Куайну, своему учителю и другу, он посвятил книгу «Опасная идея Дарвина: эволюция и смысл жизни».) В американской философии эти идеи Куайна в какой-то мере подготовили почву для *синтетических* форм философии, менее жестких и менее логицистских. Таковыми в принципе являются эволюционистские (процессуальные) философии. Деннета можно расценивать как законного наследника натурализма прагматического анализа, а его философию — как типичный вариант «американского синтеза». Сам он не противится оценкам его творчества, в которых говорится, что в итоге своей «инженерной работы» он пришел к Grand Syntheses. Свое отличие от дедуктивно и спекулятивно мыслящих философов он видит в принятии «минимальной отправной точки». Последнее означает, что мы не должны начинать с Cogito, с кантовской дихотомии субъективного и объективного, противопоставления «перволичностной» и «третьеличностной» позиций. Все объяснения должны отправляться от до-жизненного, до-сознательного уровня и идти от него вверх, нежели наоборот: от трансцендентальных принципов — божественных или иных — к природе. Иначе говоря, в его Grand Syntheses не допускаются никакие «небесные зацепки» — «только журавли в руки».

<sup>1</sup> Насколько реальна или нереальна эта «угроза», зависит от понимания научного статуса дисциплин — физики, биологии, психологии, а также статуса логики и интенциональности в этих дисциплинах. По этому поводу существуют различные мнения, но мы не будем касаться их.

245

## VII. Об условиях личностного

Сущность человека не есть абстракт, присущий отдельному индивиду. В своей действительности она есть совокупность всех общественных отношений.

*Карл Маркс*

### 1. Необходимые и достаточные условия понятия «личностное»

Смысл понятия «личность», считавшийся после Канта более или менее определенным (вменяемое и ответственное человеческое бытие), в последние полвека существенно расширился в объеме и потерял прежние более или менее четкие границы. Предметом споров стали вопросы, которые раньше не приходили в голову. Можно ли привести необходимые и достаточные условия, позволяющие правомерно приписывать признак «личностного» тому или иному существу? Основывается ли понятие «морально личностное» на существовании «метафизически личностного»? Есть ли различие между понятиями «биологическое человеческое бытие» и «личностное бытие»? Тожественны ли понятия «биологически человеческое» и «социально человеческое»? Правомерно ли считать человеческий эмбрион или новорожденного младенца личностью? Можно ли определять личность как бытие, имеющее моральное право на жизнь? Означает ли признание определенных прав животных от-

246

ношение к ним как к личностям? Можно ли рассматривать группы, партии, корпорации в качестве коллективных личностей? Следует ли относиться к роботу как к личности и признавать за ним моральные права? (Многие из этих вопросов активно эксплуатируются в кинематографе, например, вопрос о правомерности или неправомерности морального отношения к компьютеру составляет важную сюжетную линию фильма С. Кубрика «2001».)

Ответы на эти вопросы имеют не только интеллектуальный интерес; они практически важны для этики, социологии, юриспруденции, экологии. Из-за отсутствия согласия по ним в обществе нередко возникают идейные конфликты (вспомним раскол американского общества по вопросу об абортах и моральном праве зародыша на биологическую жизнь, полемику о правах животных, о клонировании человека и др.). В зависимости от идейных и моральных ориентаций участники дискуссий по-разному толкуют «личностное».

Совершенно очевидно, что грамотный ответ на эти вопросы должен основываться не на

идеологических догмах или сентиментах, а на философско-теоретическом анализе категорий, относящихся к пониманию «личностного». В свою очередь, этот анализ напрямую зависит от понимания фундаментальных вопросов о сознании, самости, свободе воли. Далеко не у всех философов есть на этот счет своя разработанная концепция. Дэниэл Деннет — заметное исключение. Теория личностного у него опирается на функционалистскую теорию сознания, концепцию интенциональности, особую трактовку самости, эпистемологический анализ объяснительных стратегий, гетерофеноменологический метод.

Как уже говорилось в предыдущих главах, анализ категорий «сознание», «самость», «интенциональность» на основе гетерофеноменологии, то есть «с позиции третьего лица» (в отличие от апеллирующего к интроспекции «перволичностного» языка), служит у Деннета средством не только для преодоления расколов и Вселенной и знания, но и для примирения конфликтующих позиций в понимании человека. Он

247

убежден, что предложенный им анализ снимает многие противоречия детерминизма и свободы, механицизма и моральной ответственности, интерпретивизма и рационализма. В практическом отношении он важен для юриспруденции, где постоянно возникают вопросы о точном содержании понятия «личность» при определении вменяемости, правовой и моральной ответственности подсудимых. И, конечно, значим для этики и социальных наук, где возникает много путаницы при неряшливом понимании относящихся к человеку категорий.

Показательна в этом отношении его работа «Условия личностного»<sup>1</sup>, перевод которой дан в Приложении 1 настоящего издания. Ее содержанием является рассмотрение возможностей логически правомерного определения понятия «личностное», а методом — концептуальный анализ обычного представления о личностном и допускаемых при этом категориальных ошибок. Мы хотели бы обратить внимание на последнее — на характер и стилистику его аргументации, а также на стратегию соединения биологического и социального подходов.

Для понимания обстоятельств появления этой статьи важно иметь в виду, что в 50—70-е годы в англоязычной литературе публикуется большое количество работ по проблеме личности (Б. Скиннер, П. Стросон, А. Айер, Х. Франкфурт, Х. Грайс, Т. Нагель, А. Рорти и др. Ход рассуждений Деннета и выбор объектов критики понятны в контексте дискуссий с этими авторами). Основной расклад противоборствующих позиций по этой проблеме в конечном счете зависит от принятия позиции по кардинальному вопросу: существует или не существует «метафизически личностное»? Мы не будем останавливаться на спорах вокруг метафизики и со-

<sup>1</sup> Dennett D. Conditions of Personhood // D.C. Dennett. Brainstorms. Philosophical Essays on Mind and Psychology. Ch. 14. Cambridge MA, L., 1986. P. 267—299. На русском яз.: Деннет Д. Условия личностного // Приложение 1 настоящего издания. Сноски даны на это издание.

248

средоточим внимание на его анализе этого понятия. Прежде чем выносить твердый вердикт о существовании такой реалии, как «личностное», следует рассмотреть смыслы понятия «личность». В традиции европейской философии и юриспруденции, говорит Деннет, принято считать как само собой разумеющееся, что все существа, относящиеся к биологическому виду *homo sapiens*, являются личностями. При этом обычно не ставится вопрос о *логической правомерности* приписывания им этого почетного звания. В то же время на практике, в частности юридической, некоторых людей, например, младенцев и психически больных, исключают из этой категории, несмотря на их принадлежность к биологическому виду человека. В связи с этим возникает вопрос, вправе ли мы отождествлять понятие «личностное» с «биологически человеческим»? Чтобы ответить на него, биологический признак *homo sapiens* можно подвергнуть маленькому мысленному эксперименту. Хотя человеческие существа являются единственными существами, которых мы принимаем за людей, и хотя термины «человеческое» и «личность» кажутся коэкстенсивными, можно вообразить ситуацию, когда, встретившись с инопланетянами, мы будем относиться к ним как к личностям, при всем том, что по своей биологии они, скорее всего, не будут входить в класс *homo sapiens*. Конечно, при условии наличия у них атрибутов, которые мы

считаем существенными, приписывая самим себе качества личностного. Пример с данной виртуальной ситуацией, говорит Деннет, свидетельствует о логической непоследовательности отождествления понятия «личность» с биологическим человеческим видом.

Одним из первых, кто занялся анализом (и последующей ревизией) понятия «личность», был Б. Скиннер. В наделавшей много шума книге «За пределами свободы и достоинства»<sup>1</sup> он высказал крамольную мысль об устарелости традиционного смысла этого понятия. Эта мысль была встречена бурей критики, доказывающей неэлиминируемость этого понятия из

<sup>1</sup> Skinner B. *Beyond Freedom and Dignity*. N.Y. 1971.

249

нашей концептуальной схемы. Деннет говорит, что он очень внимательно изучил бихевиористские аргументы Скиннера<sup>1</sup> и пришел к выводу, что многие авторы в своем возмущении не всегда были к нему справедливы. У его позиции немало изъянов, тем не менее следует проявлять осторожность в ее общей оценке и не отбрасывать с порога рациональные зерна ряда составляющих ее положений.

К числу главных теоретических изъянов Скиннера, по мнению Деннета, относится философски-теоретическая непроработанность используемых им понятий: не проделана важная работа, которую надлежало проделать, прежде чем приступать к общим рассуждениям. Она должна состоять в концептуальном и логическом анализе этого понятия. Нужно рассмотреть *необходимые и достаточные* условия, соблюдение которых позволяет логически правомерно говорить о какой-то сущности как о «личности». (Напомним читателю, в чем состоят эти условия. В логике, например, *A* является *необходимым* условием для *B*, если *B* не может существовать без *A*; но *A* является *достаточным* условием для *B*, если *A* само по себе гарантирует существование *B*. Например, бытие мужчиной является необходимым условием для бытия мужем, а бытие мужем является достаточным условием для бытия мужчиной. В случае личности имеет место другая ситуация: бытие членом человеческого рода является необходимым условием для бытия личностью, но бытие личностью не есть достаточное условие для бытия homo sapiens.)

Если в результате анализа окажется, что мы не в состоянии привести такие условия, тогда постулат о существовании такого признака, как «личностное», следует признать неправомерным. Не исключено, замечает он, что рассмотрение этих условий покажет, что искать было нечего. «Может оказаться, скажем, что понятие личности — только свободно циркулирующее по-

<sup>1</sup> В книге «Brainstorms» у него есть глава «Skinner Skinned», посвященная рассмотрению взглядов Скиннера. По-русски ее можно перевести как «Скиннер без кожи» или «Скорняк без кожи». Игра слов: Skinner по-английски означает «скорняк».

250

четное звание, которое мы с удовлетворением прилагаем к самим себе и к другим людям в той мере, в какой нами движет дух, подогреваемый нашими эмоциями, эстетическими чувствами, соображениями политики и тому подобным, — точно так же как, скажем, понятие «модный» относится ко всем, и только к тем, кто позволяет расценивать себя как модный другими, тоже считающими самих себя модными. Действительно, быть личностью оказывается *чем-то* вроде этого, а если в ее бытии нет ничего большего, нам надлежит, если удастся, пересмотреть важную значимость, приписываемую этому понятию»<sup>1</sup>.

При обычном словоупотреблении в понятии «личность» переплетаются два смысла — метафизический и моральный. Под метафизической личностью чаще всего понимают интеллектуального, осознающего себя и чувствующего агента, наделенного свободной волей, в то время как под моральной личностью понимается вменяемый ответственный агент, имеющий как права, так и обязанности. При этом метафизический смысл, связанный с понятиями «сознательный», «активный», «наделенный свободой» воли, принимается за основания морального. Некоторые философы считают метафизическое понятие личности *достаточным* условием для морального понятия «личность»<sup>2</sup>. Однако даже если предположить, что метафизически личностное существует, говорит Деннет, нужно еще доказать, каково отношение между бытием личности в метафизическом смысле и бытием личности в моральном смысле.

Оставив в стороне спор о метафизическом смысле понятия «личность»<sup>3</sup>, в данной ра-

<sup>1</sup> Деннет Д. Условия личностного. С. 451—454.

<sup>2</sup> Деннет имеет в виду статью Хэрри Франкфурта «Свобода воли и понятие личности», в которой тот предложил ряд условий для определения личностного, которые он счел не только необходимыми, но и достаточными (Frankfurt H. Freedom of the Will and the Concept of a Person // The Journal of Philosophy. VII. Jan. 14, 1971).

<sup>3</sup> Этот смысл рассматривается в других публикациях Деннета, посвященных свободе воли, в частности в книге: Elbow Room: the Varieties of Free Will Worth Wanting. Oxford, 1984.

251

боте он сделал предметом анализа условия, позволяющие нам считать какое-либо существо моральной личностью, следовательно свободной, вменяемой и ответственной.

Разными авторами приводятся различные условия, которые Деннет свел к шести основным. Для того чтобы признать какое-либо бытие моральной личностью, обычно предполагается, что это бытие должно быть:

- 1) рациональным существом;
- 2) существом, которому свойственны состояния сознания и можно приписывать ментальные или интенциональные предикаты;
- 3) существом, бытие которого конституируется в зависимости от принятой по отношению к нему другими существами установки (stance); иначе говоря, форма обращения других существ с данным существом конституитивна для его бытия как моральной личности;
- 4) существом, способным к взаимности (reciprocity). Взаимность предполагается даже в том случае, когда имеет место отношение насильника к жертве, ибо и тогда он относится к ней, как к личности;
- 5) существом, способным к вербальной коммуникации. Это условие выводит не-человеческих животных из объема понятия личностного;
- 6) существом, способным к самосознанию. Этот признак свойствен человеческому виду и не свойствен другим видам существ. Его определяют также как ту или иную форму самосознания или само-рефлексии.

Предлагаемый Деннетом концептуальный анализ содержания этих условий предполагает решение трех задач: 1) определить зависимость этих условий друг от друга; 2) показать, в каком смысле они являются необходимыми условиями; 3) выявить трудности доказательства того, что все эти условия вместе взятые являются достаточными условиями для понимания личностного. Мы не будем воспроизводить все этапы его анализа в соответствии с поставленными задачами (читатель имеет возможность познакомиться с ними по переводу

252

статьи Деннета «Условия личностного»). Скажем только, что смысл его анализа зависимостей состоит в показе того, что первые три темы внутренне взаимозависимы: быть рациональным — значит быть интенциональным и быть объектом определенной установки. Вместе взятые, они являются необходимыми, но недостаточными условиями для показа формы взаимности. Взаимность, в свою очередь, является необходимым, но недостаточным условием наличия способности к вербальной коммуникации, представляющей собой, в свою очередь, необходимое условие для наличия особого типа сознания или самосознания.

## 2. «Личностное» в свете интенциональной установки

Специфика позиции Деннета состоит в применении к анализу «личностного» его концепции интенциональности. Анализируя формы зависимостей шести обозначенных условий, он все последующие условия рассматривает как выражение *дополнительных* уровней интенциональности. Взаимность — это вопрос о том, имеет ли какое-либо сущее интенции второго уровня, то есть можно ли ему приписывать *верования о верованиях*. Коммуникация — это вопрос об интенциях *третьего порядка*, то есть о приписывании какой-либо системе интенции такого типа: я распознаю, что она имеется в виду и что я буду реагировать на это определенным образом.



Сознание включает в себя интенции *второго порядка*, т.е. приписывание системе желания иметь желание.

Как уже говорилось в главе «Интенциональность как установка рациональности», при объяснении любой системы Деннет выделяет три главные установки. *Физическую установку*, которая основывается на знании действительного физического состояния системы, например, физико-химической структуры организма. *Установку дизайна*, опирающуюся на знание

253

системной организации объекта и паттернов его функционирования, например, нейрофизиологическую организацию мозга. *Интенциональную установку*, основывающуюся на допущении рационального поведения объекта. По Деннету, все эти установки — это только наши объяснительные стратегии или инструменты интерпретации, находящиеся — в зависимости от наших прагматических целей — в постоянной подвижке. Ошибка Скиннера состояла в том, что он пытался объяснять личностное поведение агентов, используя физическую установку и установку дизайна, отбросив интенциональную установку. Тот же изъян свойствен и бихевиоризму У. Куайна. Конечно, и Скиннер и Куайн правы в том, что объяснения с минимальной интенциональностью являются предпочтительнее, поскольку только у физической установки и установки дизайна имеется онтологический статус, интенциональная установка его не имеет. Возможно, что в будущем с расширением научного знания о функционировании нейрофизиологической организации потребность в интенциональном объяснении исчезнет, но в настоящее время, убежден Деннет, без нее не обойтись. И вот почему.

Когда по отношению к какому-либо существу мы принимаем интенциональную установку, мы получаем возможность *предсказывать* поведение системы. В случае, когда нашей целью является *моральное поведение* какой-либо сущности, мы можем принимать «личностную установку» или «моральную установку» и трактовать эту сущность как свободную, ответственную личность и видеть в ней субъект морального обязательства. То есть принятие моральной установки у Деннета аналогично принятию интенциональной установки в том отношении, что она зависит не от свойства объекта с доказуемыми и каузальными — правильными или ложными — оценками и выводами из реального положения дел, а от нашей позиции и выбора, соотношенного с целями и желаниями.

Применение интенциональной установки дает возможность приписывать системам разные уровни интенциональности. Это позволяет отнести личность к *интенциональной системе*

254

*высшего уровня*, т.е. владеющей языком и поэтому осознающей себя и свое поведение. «Высший уровень» означает, что у личности, наряду с собственными верованиями и желаниями, имеются верования в наличие желаний других людей, верования о верованиях других людей относительно их собственных желаний и т.п.

Многие авторы в определении личностного на первый план выдвигают шестое условие — самосознание, а некоторые описывают это условие как своего рода феноменологическую игру. Деннет советует здесь проявлять большую осторожность, чтобы не скатиться к субъективистской позиции и не возводить свойство самосознания в ранг не подлежащих анализу трансцендентальных свойств.

Хорошую тактику предложила Элизабет Энском. Она такова: *если меня считают ответственным за какое-то действие* (какую-то часть моего поведения при определенном описании), я должна осознавать это действие при данном описании<sup>1</sup>, поскольку только при условии осознания этого действия появляется возможность высказываться о нем, обсуждать его, включаться в игру вопросов-ответов и приводить основания для своих действий. Деннет акцентирует внимание на следующем важном условии участия в этой «игре»: только агентов, обладающих способностью *приводить разумные основания*, можно склонить или, наоборот, разубедить аргументами о характере тех или иных действий или отношений. Если кто-либо не в состоянии приводить основания, отвечать на возражения и «прислушиваться к голосу разума», его нельзя считать морально ответственным. Его вывод таков: способность к вербальной коммуникации и к

осознанию своих действий суще-

<sup>1</sup> Деннет говорит: меня могут считать ответственным за события или состояния дел, не проходивших через мое сознание, но некоторые из них я обязан был осознавать, однако они относятся к интенциональным действиям. В этих условиях я ответствен за развитие событий в силу ответственности за предсказуемые последствия действий — включая акты оплошности, которые мне известны.

255

ственна для проявляющих готовность прислушиваться к аргументу или убеждению, а такое убеждение и взаимное приспособление интересов, достигаемое при взаимной эксплуатации рациональности, является характерной чертой оптимальной модели личностного взаимодействия.

В этих случаях *доступ к себе* как к личности такой же самый, как и установка по отношению к другим. Человек-личность должен обладать способностью *спрашивать себя*, каковы в действительности его собственные желания, мотивы, основания, и только если он в состоянии осознанно ответить на эти вопросы, осознать свои мотивы и желания, он обретает позицию, позволяющую *заставить изменить себя*. В этом состоит сущность рационального в сфере морального.

Подводя итог проделанному анализу понятия «личностное», Деннет тем не менее выражает скепсис относительно того, что перечисленные шесть условий личностного могут быть приняты за твердое основание определения понятия «личность». Их рассмотрение показало, что все они являются *необходимыми условиями*, однако трудно сказать, являются ли они все вместе взятые достаточными условиями для морально личностного. Причем, замечает он, «допущение, что некое существо является личностью, оказывается шатким как раз в тех случаях, где в практическом отношении оно особенно значимо — при совершении каких-то проступков и возникновении вопроса об ответственности»<sup>1</sup>. Почему же нам не удастся найти твердое основание для определения «личностного»? Потому, отвечает Деннет, что данное понятие, являясь предметом интенциональной установки, неизбежно *нормативно*. Иначе говоря, решение включать или не включать то или иное бытие (человека, животных или компьютеры) в сообщество личностей является произвольным и по большей части прагматическим.

Из-за невозможности приведения удовлетворительных достаточных условий всякая попытка определения какого-либо

<sup>1</sup> Деннет Д. Условия личностного. Р. 477.

256

существа как личности всегда может быть подвергнута сомнению. Все вышеперечисленные шесть пунктов являются *необходимыми* условиями, но их нельзя признать *достаточными*. Они не гарантируют, что какое-либо живое существо является личностью. Из-за невозможности приведения объективно удовлетворительных достаточных условий всякая попытка определения какого-либо существа как личности оказывается шаткой. Вывод о том, что субстанции «морально личностного» не существует, должна учитывать юриспруденция, в которой решение вопроса о правовой и моральной ответственности упирается в вопрос, кого мы можем считать личностью. Общий вывод Деннета таков: Маркс и Ницше пошли по верному пути, когда приписывали сознанию коммуникативно-социальный характер, а личности — нормативный характер, хотя и не представили логически корректной аргументации.

### 3. Нормативно личностное и метафизически личностное

Несмотря на позицию нормативизма в определении личностного, мы не найдем у Деннета жесткого элиминативистского вердикта в отношении «метафизически личностного». Остается неясность в его оценке отношения «морально личностного» и «метафизически личностного». С одной стороны, у него есть заявления о том, что он не исключает существование «метафизически личностного» (возможно, замечает он, это даже верно), которое отличается от «морально личностного» и лежит в основе последнего. Если такое вдруг обнаружится, тогда «по-видимому, должно существовать основание для предположения, что метафизически личностное является

необходимым условием морально личностного»<sup>1</sup>. Отсутствие твердого элиминивистского вердикта он связывает с тем, что

<sup>1</sup> Там же. С. 451.

257

трудно отрицать существование, по меньшей мере, одного бытия, относительно которого принята к нему установка не является делом прагматического выбора. Это я сам. «Я не могу устраниваться от картины самого себя как личности»<sup>1</sup>. «Я должен рассматривать самого себя как личность»<sup>2</sup>. «Я — личность, и вы тоже являетесь ею. По большей части это заявление не вызывает сомнения»<sup>3</sup>. Метафизическая личность в этом случае выступает как «сущность, которой приписываются состояния сознания или само-сознания»<sup>4</sup>. С другой стороны, в противоречие этому заявлению он утверждает, что при определенных условиях «Мы не можем даже в отношении нас самих сказать, являемся ли мы личностями»<sup>5</sup>. Ибо только определенное (моральное) обращение с нами самими со стороны других людей конституитивно для нашего бытия как личности. Иначе говоря, наше личностное бытие зависит от нашей собственной интенциональной установки и интенциональных установок других людей.

Критики упрекают Деннета за то, что представленные им условия для определения личностного с помощью идеи интенциональности и нормативизма одновременно являются и слишком узкими, и слишком широкими. Например, когда мы имеем дело с душевнобольным человеком, мы не вправе, если исходить из концепции Деннета, приписывать ему предикат морально личностного и считать ответственным. Но это слишком узкое понимание нормативизма. Согласно существующим в обществе нормам, к душевнобольным относятся как к личностям, исходя из идеальной посылки, что в случае излечения их можно признать вменяемыми и возратить юридические права. То же самое можно сказать о маленьких детях (или человеческих эмбрионах). К поведению младенца, конечно,

<sup>1</sup> Dennett D. *Brainstorms*. P. 254.

<sup>2</sup> Ibid. P. 255.

<sup>3</sup> Ibid. P. 267.

<sup>4</sup> Ibid. P. 268.

<sup>5</sup> Ibid. P. 285.

258

бессмысленно применять интенциональную установку, предполагающую обмен рациональными аргументами. Тем не менее дети наделены правами на все необходимое для поддержания их жизни, хотя права временно делегированы взрослым, берущим на себя моральные обязательства. Ограниченность деннетовского толкования нормативизма видна и на примере отношения общества к умирающему человеку. Все люди имеют право на достойную смерть именно как личности, хотя не могут выступать объектами интенциональной установки, т.е. отвечать взаимностью, вступать в коммуникацию и т.д. Даже с полным идиотом, не осознающим своих намерений (совсем не отвечающим требованиям Деннета), общество обращается как с личностью. Стандартное возражение Деннета на критику и приводимые им контрпримеры состоят в том, что в нашей объяснительной стратегии, в частности для целей юриспруденции, имеет смысл варьировать установки. По отношению к безумцам или новорожденным младенцам, не владеющим языком, прагматически целесообразными и предсказательными окажутся физическая установка или установка дизайнера, нежели интенциональная установка.

Одновременно, считают критики, Деннет слишком широко толкует интенциональность. В предложенной им схеме у компьютера больше шансов получить почетное звание «личность», нежели у людей, не отвечающих принципу интенциональности и рациональности. При соответствующих обстоятельствах (например, при игре в шахматы) компьютеру вменяется интенциональность и рациональность. Компьютер у него является в этих случаях интенциональной системой, потому что он поддается определенной установке, принятой в отношении него, а именно интенциональной установке. В то же время запрещается приписывать это почетное звание людям, которые не соответствуют его представлению о нормах рациональности.

Один из критиков — Д. Уилсон — пишет, что если следовать логике рассуждения Деннета, придет время, когда компьютер, если у него будет скорректированное понимание рациональности, сможет отрицать наличие морально личностного

259

у человека. Если предположить, что цель Деннета (и Куайна) в отношении психологии когда-либо будет реализована, то есть настанет час, когда человеческое бытие целиком и полностью будет объяснено с точки зрения установки дизайна, то произойдет это скорее всего с помощью компьютера. Допустим, что удалось спроектировать компьютер, способный выполнить такую задачу: он будет настолько софистичной и способной к самообучению машиной, что с ним невозможно будет общаться, не принимая ее за бытие, способное к саморефлексии, взаимности, коммуникации и т.д. В этом случае не только людей, но и компьютеры нужно будет считать моральными агентами. Из этой ситуации напрашивается вывод, что, с точки зрения компьютера, было бы неверно видеть в людях моральных агентов, ибо компьютер объясняет и предсказывает их поведение, имея в виду только дизайн человека. Поэтому, даже признав права компьютера, следует допустить, что с позиции компьютера вполне правомерно отрицать наличие таких прав у людей. Свою критику Уилсон заканчивает вопросом: «Можно ли считать правильным объяснение моральной обязанности, ведущее к таким выводам?»<sup>1</sup>. Причину неудач Деннета он видит не только в том, что понимание морального он поставил в зависимость от условий, которые по большей части иррелевантны к морали; Дж. Роулс сделал то же самое, но получил хорошие результаты. Дело в том, что Деннет рисует моральные позиции в состоянии постоянной текучести и лабильности: приписывание моральных качеств той или иной системе зависит только от наших прагматических целей и умений предлагать объяснения с минимальной интенциональностью.

<sup>1</sup> Wilson D.C. Functionalism and Moral Personhood: One view considered // Philosophy and Phenomenological Research. V. XLIV. № 4. June 1984. P. 529.

## **VIII. О ПОНЯТИИ «ОТВЕТСТВЕННОСТЬ» В СВЕТЕ МЕХАНИЦИСТСКОГО ОБЪЯСНЕНИЯ ЧЕЛОВЕКА**

Наше равнение на совершенную способность практического разума Канта, возможно, является столь самонадеянным, что гордая самоидентификация самих себя в качестве моральных агентов есть мания величия.

*Дэниел Деннет*

### **1. Можно ли включить целевое объяснение человека в физикалистскую картину мира?**

Для материалистической философии с ее интенцией к унификации знания объяснение феномена целеполагающего и ответственного поведения личности является предметом давнего спора. Со времен Ламетри, а может быть и раньше, в ней идет дискуссия о совместимости или несовместимости детерминизма (или индетерминизма) с моральной ответственностью или, в более современной формулировке, о соотношении механицистски-каузального и целевого объяснений человека. Сегодня острота споров усугубляется на фоне интенсивного развития нейронаук фактом, что в сферу механицистского объяснения попадает не только тело, но и мозг; в принципе возможны имплантация или удаление ткани нейронов и существенное изменение характеристик личности. Уже не выглядит фантастикой коррекция

261

патологического поведения людей с помощью таких операций. Не говоря уже о получившей развитие технике манипуляции людьми при помощи массового гипноза, средствами масс-медиа или действующей на подсознание рекламой: объекты манипуляции могут совершать поступки, цели которых замышляют другие люди. В связи с этим возникает масса философских вопросов, решение которых имеет прикладную значимость в медицине, в моральных теориях, в юриспруденции, в частности для определения вменяемости и ответственности человека.

В нынешней англоязычной философии эта проблема чаще всего ставится не столько в форме отношения детерминизма (или индетерминизма) и ответственности, сколько в более общем виде как отношение *механицистских объяснений и объяснений, основанных на целеполагании и ответственности*. Человек убежден, что, в отличие, скажем, от кирпича или животного, ему присуще особое свойство — сознательное целеполагание и моральная ответственность за действия. В то же время он не может не заметить тенденции, что по мере увеличения доли механицистских объяснений его поведенческих проявлений целевое объяснение сужает свой диапазон. Простой пример: удар, нанесенный одним человеком другому, пострадавший может оценивать как целенаправленный акт, за который ударявший должен нести ответственность; объяснения доктора, что у ударявшего был приступ эпилепсии, снимают ответственность с него и делают целевое объяснение излишним. Другой пример: калькулятор, несомненно, значительно быстрее (и точнее) осуществляет математические операции. В отличие от него ребенок, складывающий 2+2 и получающий 4, делает это целенаправленно. Помимо логических ходов в его действиях присутствует «добавка» в виде цели и ответственности (получить похвалу учителя, удовольствие от успеха, желание выполнить задание и др.). Из сопоставления этих примеров возникает вопрос, возможно или невозможно соединить в единую картину механицистское объяснение и

262

личностные «добавки» в виде целеполагания и ответственности.

В аналитической философии сложилось несколько позиций относительно совместимости или несовместимости механицизма и ответственности — радикальные и умеренные. В других работах мы уже писали о них<sup>1</sup>.

Одной из радикальных является позиция Пола Черчленда, которую иногда именуют «агрессивным сциентизмом», «физикалистским империализмом» или, более мягко, «обещающим материализмом». Согласно ей, будущая совершенная (идеальная) наука в принципе сможет механицистски объяснить всю «фабрику человека», хотя на практике это вряд ли достижимо. К «соглашательской» можно отнести позицию Уилфреда Селларса. Конфликт механицистски-объективистского и личностно-целеполагающего объяснений, или в его терминологии «научного» и «наличного» образов мира («Scientific Image» and «Manifest Image»), он оценил как «трагичный» для научного мировоззрения. Суть трагедии состоит в нередуцируемости «должного» к «сущему» (ought to is). Отличительная особенность личности состоит в ее следовании нормам и ответственности. Установка же физикализма предполагает, что все эти категории рано или поздно будут объяснены в физикалистских терминах. Селларс не исключал, что в будущей научно-материалистической картине мира с единой системой отсчета научное изображение человека, возможно, окажется изображением сложной физической системы. Все же ему представлялось, что изображение личности только в физических терминах нереалистично; «научный образ» не может заменить «наличный образ», не отвергая его собственные основания. «Наличный образ выступает каркасом, в котором отличительные черты личностей концептуально являются несводимыми к чертам неличностей, т.е.

<sup>1</sup> См.: Юлина Н.С. Проблема человека в философии физикализма // Буржуазная философская антропология XX века. М., 1986. С. 133—159; О позиции П. Черчленда см.: Юлина Н.С. Очерки по философии в США. XX век. М., 1999. С. 180—197.

263

животным и просто материальным вещам»<sup>1</sup>. В конечном итоге Селларс склонился к мысли, что концептуальный каркас, в рамках которого мы описываем личность («Manifest Image»), вряд ли можно *примирить* с научным, объективистским объяснением мира («Scientific Image»). Скорее он таков, что может быть только *присоединен* к нему.

Позиция «присоединения», а не «включения» категории личности в монистическую физикалистскую картину мира свидетельствует об ее аномальности. Перед физикалистами постоянно возникает альтернатива: либо признать, что личности являются сложными материальными телами, поведение которых полностью объяснимо в рамках механицистского



панобъективизма, либо признать, что присущие им интенциональность, целеполагание, свобода, ответственность, самотворчество не укладываются в эти рамки. В случае признания последнего следует констатировать, что либо физикалистский материализм имеет аномалии («аномальный монизм» Д. Дэвидсона), либо следует согласиться, что философия, пользующаяся менталистскими и личностными категориями, не исчерпала своих возможностей. Далеко не все философы аналитического склада готовы пойти на попятную и допустить принципиальную нередуцируемость должного к сущему. Большинство из них пытаются найти новую аргументацию в пользу физикализма «с человеческим лицом». Одни ищут лингвистические и логические ловушки в оппозиции механицизма и ответственности, другие выстраивают дополнительные теоретические конструкции.

Деннет пошел по последнему пути. Разбираемая нами работа «Механицизм и ответственность» (она включена в книгу «Головоломки», 1986)<sup>2</sup> интересна с точки зрения обо-

<sup>1</sup> Sellars W. *Science, Perception and Reality*. L., 1963. P. 40; Sellars W. *Fatalism and Determinism // Freedom and Determinism*. Ed. by K. Lerer. N.Y., 1966. P. 145.

<sup>2</sup> Dennett D. *Mechanism and Responsibility // Dennett D. Brainstorms*. Ch. 12. Cambridge MA, L., 1986.

264

снования тезиса о *совместимости механицизма и ответственности* с помощью применения теории интенциональности. Безотносительно к тому, насколько успешной оказалась эта попытка, данная работа показательна своей идеологической интенцией — развеять бытующий негативный образ физикализма и опасения, что эта философия подрывает сложившийся в западной культуре образ человека как существа морального и ответственного.

Как и большая часть из написанного им, данная работа Деннета полемична. Ее логика в полной мере понятна в контексте его комментариев позиций П. Стросона, А. МакКея, Н. Малколма, А. Макинтайра, Э. Флю, У. Селларса и других. Мы не будем касаться всех затронутых Деннетом тем и воспроизводить полемику с оппонентами. Наша задача сводится к экспликации его позиции по следующим вопросам: применение теории установок для обоснования тезиса о совместимости механицизма и ответственности; взаимоотношения интенционального, рационального и морально личностного; действие и ответственность; манипуляция и ответственность.

## 2. Совместимость механицизма и ответственности в свете теории установок

Первым шагом к примирению механицизма и ответственности, по Деннету, должно быть прояснение природы этих конфликтующих типов объяснения. В случае механистического объяснения в деятельности человека фиксируются физические процессы, химические реакции, электрические импульсы и т.п., и на их основе дается каузальное (основанное на законах науки) объективное объяснение. Объяснения, служащие для обоснования вердикта ответственности, выражены в терминах верований, интенций, желаний и мотивов поведения личности. (Это то, что можно отнести к кантовской сфере

265

практического разума.) Они именуются «целевыми», «рациональными», «деятельностными» или «интенциональными» (последний термин предпочтительнее). Считается, что, в отличие от механистических объяснений, интенциональные объяснения не являются каузальными.

Расширение сферы механицистского объяснения человека побуждает некоторых философов заявлять о его приоритетности и о последующем вытеснении целеполагающего, что предполагает ложность любого объяснения в терминах желаний, верований, интенций и ответственности. Джон Хосперс, например, считает, что «чем глубже и детальнее мы познаем каузальные факторы, вынуждающие личность вести себя так, как она ведет себя, тем больше мы склоняемся к освобождению ее от ответственности». Согласно Норману Малколму, «хотя целевые объяснения не могут быть зависимыми от не-целевых объяснений, они, скорее всего, будут отвергнуты на основе верификации всеобъемлющей нейрофизиологической теории поведения»<sup>1</sup>. Очень многие философы не согласны с ними, полагая, что редукция должного к сущему невозможна. Деннет

считает возможным подойти к решению конфликта более гибко. Заявлять категорически, что для таких утверждений вовсе нет оснований, говорит он, было бы преждевременно и неострожно. В аргументах Хосперса и Малколма есть изъяны, но они устранимы, что делает приемлемым другой тезис — о *совместимости* механицизма и ответственности.

Для устранения изъянов Деннет предлагает применить к отношению механицизма и ответственности различные установки или объяснительные стратегии, а именно — физическую установку, установку дизайнера и интенциональную установку, а затем посмотреть, как они взаимодействуют.

<sup>1</sup> Hospers J. What means this Freedom? // *Determinism and Freedom in the Age of Modern Science*. Ed. by S.Hook. N.Y., 1958. P. 133; Malcolm N. The Conceivability of Mechanism // *Philosophical Review*. LXXVII. 1968. P. 51. Цит. по: Dennett D. *Brainstorms*. 1986. P. 234.

266

Объяснения с точки зрения первых двух установок — дизайнера и физической — являются механистическими и каузальными. Интенциональные объяснения не являются таковыми: от них нельзя требовать того, что могут дать только первые две установки. Часто интенциональность путают с каузальностью. Однако далеко не все содержащиеся интенциональные термины объяснения являются интенциональными. Нередко верование или желание рассматриваются как причина или (редко) как следствие в юмовском смысле причины и следствия. Например, высказывание «Его вера в то, что ружье заряжено, стала причиной его сердечного приступа» кажется интенциональным, на самом деле оно является каузальным гибридом. Здесь соединяются два события и нет логического основания для explicandum. Подлинно интенциональным оно было бы в случае такой формулировки: «Он бросился на пол, поскольку верил, что ружье заряжено».

Преимущественная сфера интенциональных объяснений — действия и поведение личностей. Главная их нормативная особенность — *допущение рациональности* в объясняемой системе, то есть допущение ее разумности в свете приписываемых ей верований, интенций, желаний и целей, которые соответствуют ее интересам. Особенность механистических объяснений состоит в том, что они исходят из *не-рациональности* изучаемых систем (но не их иррациональности). Приписываемая поведенческим системам рациональность не обязательно должна быть совершенной. Отличительный признак интенциональной установки — терпимое принятие рациональности. Когда мы общаемся с другими взрослыми людьми, мы начинаем с допущения их рациональности и затем приспособляем предсказания их поведения по мере накопления нашего опыта общения. Когда они начинают вести себя иррационально, мы обращаемся к другим объяснительным стратегиям, прежде всего к установке дизайнера, и ищем объяснения в телесных или функциональных неполадках организма. Это происходит, например, при общении с психически нездоровыми людьми. Что касается физической установки, то ее приме-

267

нение к поведенческим системам затруднительно. В сфере поведения (в отличие от деятельности внутренних органов) вряд ли когда-либо будет получено достаточное знание о нейрофизиологических процессах человека для эффективного применения физической установки (механистического объяснения), за исключением разве немногих драматических сфер, вроде хирургического лечения эпилептических припадков.

Эта особенность интенциональной установки — посылка рациональности — лучше всего видна, если применить ее к играющему в шахматы компьютеру, являющемуся механической системой. Мы предсказываем его реакцию на движения фигур, рассчитывая последующие разумные шаги. Партнер, начинающий игру, исходя из посылки иррациональности компьютера, принял бы крайне рискованную тактику. Данный пример предполагает не только отсутствие неполадок в компьютере, но также рациональность дизайнера и программирования. Конечно, интенциональная установка бессмысленна, когда мы не верим в рациональность системы, например, в предсказаниях погоды не имеет смысла задаваться вопросом о следующем «умном» движении Мудрого Старого Ветра.

Может показаться, что если партнером компьютера является его проектировщик, то он может

прибегать к установке дизайна и извлекать для себя выгоду из известных ему погрешностей рациональности, присущих программе данного компьютера. На ранних этапах проектирования шахматных программ эта тактика была выполнима, однако с созданием программ, способных к самокорректировке, его создатели уже не в состоянии придерживаться установки дизайна в игре против их собственных программ. Для того чтобы перехитрить машины, они вынуждены прибегать к интенциональной установке. Еще меньше у них шансов перехитрить машину при применении физической установки. Как и в случае объяснения поведения людей, реакции компьютера на ввод информации невозможно объяснить чисто физическими процессами, например, движением электрической энергии от клавиш до принтера или свойствами используемых материалов.

268

Деннет оговаривает, что он часто прибегает к примеру играющего в шахматы компьютера, поскольку на нем лучше демонстрировать различие объяснительных установок. Этим он вовсе не хочет сказать, что люди в принципе не отличаются от компьютеров. В сравнении с человеком самый софистичный компьютер, конечно, является в высшей степени *ограниченной системой*. Стоит только предложить ему сыграть партию в бридж, и он окажется простой железкой. В принципе можно спроектировать компьютерную систему, способную адекватно реагировать (в соответствии с некоторыми обусловленными целями) на какое-либо ограниченное число фиксированных вводов информации или конечное число переменных «значимостей» окружающих условий, однако «не существует пути к проектированию дизайна системы, с гарантией адекватно реагирующей при *всех* условиях окружающей среды»<sup>1</sup>. Исходя из факта, что мы с легкостью переходим от одной игры к другой, многие люди сделали вывод, что, в отличие от компьютера, человек является системой, способной адекватно реагировать на ввод информации *любого типа*. Действительно, когнитивные, информационные и адаптивные возможности человека по сравнению с компьютером огромны, однако при этом не следует забывать, что и человек спроектирован природой *избирательно* реагировать на окружающую среду. Многое имеющееся в его информационном поле он не воспринимает. Поэтому данный параметр различия не является абсолютным.

Вывод Деннета таков: поскольку мы можем применять интенциональную установку по отношению к механическим системам, а механическую установку в отношении людей, утверждение об их антагонизме является ложным. Иначе говоря, эти разные объяснительные стратегии вполне совместимы.

К тактике применения различия установок к личностям прибегают и другие авторы, например, А. МакКей, П. Стросон, А. Макинтайр, Х. Патнэм и др. На первый взгляд их установки кажутся сходными с моими, говорит Деннет. А. Мак-

<sup>1</sup> Dennett D. *Brainstorms*. P. 244.

269

Кей, например, проводит различие между «личностным аспектом» и «механическим аспектом» некоторых систем, оговаривая, что выбор установки заключен «внутри нас»: он есть вопрос нашего решения, а не открытия, то есть не является делом доказательств. Сходная позиция у Питера Стросона, когда он противопоставляет «объективное отношение» к чему-либо «участию в человеческих взаимоотношениях». Стремясь сделать выводы для этики, вытекающие из различия «объективного» и «личностного», эти авторы, считает Деннет, заранее наделяют личностное моральным измерением и тем самым допускают ошибку. Обсуждая следствия, вытекающие из нашего принятия отношения к какому-нибудь человеческому существу, МакКей говорит: «На личностном уровне Джо предъявил некоторые личностные требования к нам, а мы — к Джо. У нас нет возможности непосредственно вторгаться в его мозг или свободно копаться в его теле... Он стал «одним из нас», членом лингвистического сообщества и, заметьте, не в силу особого вида вещества, из которого состоит его мозг, ... а благодаря особому роду взаимной интеракции, которая может поддерживаться в сопряжении с нашей интеракцией и которую на личностном уровне мы описываем как личность-к-личности»<sup>1</sup>.

Ошибку МакКея Деннет видит в том, что два выбора он представил в виде одного. Первый выбор — восхождение от механистической к интенциональной установке не имеет морального измерения (это было показано на примере проектировщика, играющего в шахматы с компьютером). Человека нельзя обвинить в моральном уродстве, если он, находясь в отношении интеракции с компьютером, лишает его членства в своем сообществе и, соответственно, моральной ответственности. Второй выбор — принятие подлинно моральной установки по отношению к системе (и таким образом рассмотрение ее как личности) в логическом отношении — другого порядка. Моральность опирается на интенциональность, но не тождественна ей.

<sup>1</sup> Цит. по: Dennett D. *Brainstorms*. P. 240.

270

Много путаницы возникает из-за смешения морально личностного с интенциональным. При поверхностном рассмотрении личностная установка кажется простым присоединением морального обязательства к интенциональной установке. Чтобы прояснить ситуацию, Деннет помимо трех вышеназванных установок выделяет четвертую — *личностную или моральную установку*. Необходимым условием личностной установки является наличие интенциональной установки. Первая надстраивается над второй, но не тождественна ей. Интенциональная установка не предполагает никакой иной, находящейся ниже ее, установки. У нее более широкий спектр действия: она применима к растениям, животным, компьютерам, социальным организациям и другим поведенческим системам. У личностной спектр ограничен. Мы можем находиться в интенциональном, но не в моральном отношении с компьютером. Моральная установка предполагает коммуникацию и взаимность с использованием особых инструментов — языка, семиотических средств и культуры. Только они дают возможность интенциональной системе превращаться в саморефлективную, самооценивающую себя систему, способную нести ответственность за цели и действия. Поэтому нельзя редуцировать моральное к интенциональному, хотя без второго невозможно первое.

Еще одна, часто совершаемая ошибка состоит в том, что в смысл интенционального авторы вкладывают «человеческое», «эмоциональное», «заботливое», «коммуникативное», «интерактивное», а в механическое — «холодно рациональное», «расчетливое», «бездушное». По мнению Деннета, они представляют и то и другое в ложном свете. Интенциональная установка охотника к его добыче может быть сколь угодно рационально холодной и расчетливой, и, наоборот, кто-то может испытывать к его автомобилю море сантиментов. «Этический вывод, который может быть извлечен из различия установок, состоит не в том, что интенциональная установка есть некая моральная установка, а в том, что она является предусловием любой моральной установки, и отсюда, если в силу какого-нибудь триумфа механицизма она подвергается риску, вместе с ней подвергает-

271

ся риску и понятие моральной ответственности»<sup>1</sup>. Общий вывод Деннета таков: отличительным признаком интенциональной установки по отношению к личностным системам является не существование метафизически личностного, не интеракция «лицом к лицу» и не просто рациональное или моральное к ним отношение, а предположение о *разумности* этих систем. «Мы не просто разумно рассуждаем относительно того, что будут делать интенциональные системы, *мы разумно рассуждаем о том, как они будут разумно рассуждать*»<sup>2</sup> (выделено мною — Н. Ю.). Для предсказания поведения это предположение важнее, нежели накопленные факты обращения с этой системой и индуктивные подсказки о возможных ее действиях. Когда говорят о гегемонии механистического объяснения, ее неверно толковать как вытеснение интенционального. Дело обстоит по-другому. Механистические предсказания, которые не прибегают к посылке о рациональности исследуемых систем, могут опровергать интенциональные предсказания в случае, когда в реакциях системы проявляется дефицит рациональности или слабость ее «дизайна», ведущая к физически предсказываемому разрушению. В случае нормальности систем без интенциональной установки не обойтись.

### 3. Действия, манипуляция и ответственность: каузальные и рациональные факторы

Много путаницы Деннет видит вокруг другой, связанной с первой, проблемы — отношения действий и ответственности. В литературе распространена точка зрения, что человек ответствен только за события, которые являются следствиями его интенциональных *действий*, и за результаты, которые он был

<sup>1</sup> Dennett D. *Brainstorms*. P. 242—243.

<sup>2</sup> *Ibid.* P. 243.

272

в состоянии предвидеть. Много написано по поводу различия действий от простых событийных обстоятельств. При этом не обращено должное внимание на рациональные основания действий. Акт действия является исключительной привилегией рациональных существ — личностей, поэтому утверждение, что имело место действие, должно основываться не столько на исследовании предшествующих каузальных факторов, сколько на наблюдении, уместен ли в данном контексте разговор *о разумных основаниях* для действий личностей. Поэтому обычно исключают нездравомыслящих людей, с которыми невозможно вести разумный разговор; исключается и воздействие физических форс-мажорных факторов, против которых разум бессилён, являются ли эти факторы внешними (стечение обстоятельств) или внутренними (боль, заставившая меня закричать).

Согласно Деннету, мы можем говорить о наличии разумного основания для действия в том случае, когда против него можно *выдвинуть возражение*. Разумное основание можно оспаривать, приводя доводы, с некоторой надеждой на успех. Оспаривать же каузальную цепь событий, приведшую к тому или иному действию, невозможно; нельзя, например, спорить с глухим. Полезное различие между приведением разумных оснований для действий и приведением причин для действий тем не менее часто является источником еще одной вводящей в заблуждение интуиции относительно существования антагонизма между механизмом и ответственностью. Считается, что презентация аргументов, то есть мысленных действий, не может влиять на цепь причин. Однако высказываемые аргументы имеют все виды влияния на каузальную среду: их произнесение приводит в движение волны воздуха, причинно действует на вибрацию слуховых перепонки и оказывает трудно идентифицируемые наукой, но важные воздействия на мозг слушателей. (Напомним, что в понимании происхождения сознания Деннет исходит из гипотезы о влиянии языка и меме-вирусов на органику мозга.) Хотя аргументативная деятельность не оказывает зримого эффекта на траекторию полета

273

пушечного ядра или, что нам ближе, на автоматическую нервную систему, дизайн перцептивной системы мозга таков, что он чувствителен к видам передачи энергии, наличие которой следует предположить, чтобы рационально объяснить коммуникативную силу аргумента.

Много концептуальных загадок в старой проблеме механицизма и ответственности, говорит Деннет, возникает в связи с распространением новых форм *манипуляций людьми*. Случаи манипуляции, подчеркивает он, всегда требуют конкретного анализа: есть простые и сложные, подлинные и мнимые случаи манипуляции. Научные фантасты любят обыгрывать идею о вторжении нейрохирурга в мозг с целью «переписать» его и внедрить в него отсутствовавшие до этого верования или желания. У этой темы есть интересные вариации, далеко не все из которых являются фантастическими. Например, сумасшедший ученый мог бы хорошо изучить конституцию человека (или программу) и, исходя из установки дизайнера, решить, что ввод в нее определенной информации может репрограммировать человека, безотносительно к его рациональному выбору. Скажем, заставить поверить, что Мао является Богом. На практике бывают и случаи «взлома» сознания личности в гипнозе, и случаи «промывки мозгов», поэтому вопрос об ответственности здесь не является только академическим. Некоторые формы психотерапии с применением лекарств тоже подпадают под эту рубрику. Тем не менее, если



непосредственным воздействием на мозг человека заставляют поверить, что Мао является Богом, возникает много загадок и неясностей, связанных с ответственностью.

Некоторые философы полагают, что, если введением электрода в мозг мы причинно (механически) заставили человека поверить, будто он Наполеон, это снимает с него ответственность за последующие действия. Такой вывод Деннету не кажется самоочевидным. Получается, что, если обычный легковверный человек начинает верить в какой-нибудь абсурд, ему не прощают действий, совершаемых в соответствии с этим абсурдом, а если абсурдность индуцирована в его голову с помощью

274

электрода, это делает его действия безответственными. Предположим другую ситуацию: вполне благонамеренный нейрохирург имплантирует в головы особо опасных преступников убеждение, что наилучшей формой поведения для них является честность; следует ли из этого, учитывая механическое (нерациональное) воздействие хирурга, что этих людей следует лишиться статуса ответственных агентов общества? «Мне представляется, — говорит Деннет, — что нерациональность не является тем, что должно приписываться *содержанию* верований агента, но каким-то образом манере, в какой ему верится в них или как они приобретаются»<sup>1</sup>. Введение злого или доброго манипулятора ничего не проясняет в проблеме ответственности. Если мы допустим, что наличие только одного не-рационально (физически) индуцированного верования освобождает человека от ответственности, и если абсурдность или приемлемость верования не зависят от того, приобретены ли они рационально или не-рационально, мы никогда не будем уверены, является ли человек ответственным за свои действия.

Имея это в виду, следует пересмотреть случаи с преступниками, совершившими тяжкие преступления и подвергнувшимися хирургическому вмешательству с целью личного выздоровления. Предположим, что все они стали абсолютно честными. Являются ли они в этом случае ответственными гражданами или просто зомби? Если работа хирурга была тонкой и их рациональность улучшилась, преступников скорее всего следует признать ответственными. Однако в этих случаях хирург, скорее всего, будет имплантировать не столько какое-либо убеждение, сколько систему внушений, устраняющую барьеры предрассудков и ложных установок таким образом, чтобы преступники поверили во внушаемое содержание при наличии соответствующих видов доказательной и разумной поддержки.

Вывод Деннета таков: напрямую и физически нельзя имплантировать убеждение. Убеждение относится к тому, что

<sup>1</sup> Dennett D. *Brainstorms*. P. 249.

275

должно быть *одобрено* агентом (путем включения в его систему мысли или исключения ее) на основе его разумного рассуждения и приспособления ко всем остальным его верованиям. «Можно создать зомби либо хирургически, либо при помощи промывки мозгов, можно было бы даже внедрить в человека большую систему ложных верований. Тем не менее, даже если все это проделать, их настойчивое присутствие именно как *верований* будет зависеть не от крепости каких-нибудь хирургических швов, а от их способности выиграть соревнование в конфликтующих претензиях на доказательное раскрытие собственных интенций»<sup>1</sup>. Все зависит от того, насколько они встраиваются в систему рациональных верований личности и логически релевантных соображений, и в какой мере они соответствуют обстоятельствам жизни.

Рассмотрев ряд реальных и виртуальных случаев с вторжением механистических факторов в рассуждающую деятельность индивида, Деннет делает следующий обобщающий вывод: все они свидетельствуют о том, что «интенциональная установка по отношению к человеческим существам, являющаяся предварительным условием любого приписывания кому-либо ответственности, может сосуществовать с механистическим объяснением их движений»<sup>2</sup>.

Данный вывод можно оценить таким образом, что в общем и целом Деннет воспроизводит компромиссную позицию «присоединения» механицизма и целеполагания. Однако по большому счету она его не устраивает, как не устраивала бы она его учителя Куайна. Он не исключает правоту

физикализма Куайна и возможность принятия (в идеале) механистической (объективистской) установки не только по отношению к человеческим телам, но и к поведенческим актам. При этом остается без прямого ответа важный для физикалистского мировоззрения вопрос: можно ли полностью отбросить интен-

<sup>1</sup> Dennett D. *Brainstorms*. P. 252—253.

<sup>2</sup> *Ibid.* P. 253.

276

циональную установку — концептуальные области морали, агентов и ответственности — и принять когда-нибудь в далеком будущем чисто механистическое мировоззрение, или эту альтернативу следует исключить по логическим и концептуальным соображениям?

П. Стросон, С. Хемпшайр, У. Селларс и многие другие авторы в общем и целом сходятся в выводе о логической невозможности рассмотрения человека исключительно объективистским (механицистским) способом. Один из серьезных аргументов состоит в том, что утверждение об элиминируемости интенционального подразумевает существование, по меньшей мере, одной личности, к которой интенционально адресуется это утверждение, — нас самих. По отношению к самому себе объективистская позиция по логическим соображениям не может быть единственным видом позиции, которую следует принять. Деннет призывает к осторожности в отношении того, что может быть принято, а что не может. Из тезиса, что всеохватывающая механицистская теория не может быть истинной по отношению к человеческим существам, вовсе не следует, что механицистская объективистская методология к ним вообще неприменима. Приведенные им аргументы он считает достаточно убедительными, что говорить о несовместимости интенционального и механистического объяснений нельзя. Из данного тезиса не следует, что мы всегда будем характеризовать некоторые вещи интенционально, ибо «все мы можем в следующую неделю превратиться в зомби, или каким-то иным путем человеческий вид может лишиться способности к коммуникации и рациональности»<sup>1</sup>. Единственное, что можно сказать по этому поводу: будучи *личностями*, мы не можем совершенно элиминировать интенциональную установку.

Со времен Ламетри одной из опасностей, маячащей на горизонте механицизма, является фатализм. По Деннету, опасность преодолима, если исходить из неопределенности

<sup>1</sup> *Ibid.* P. 254.

277

и непредсказуемости как физического мира, так и действий личностей. Интенциональные системы наделены эпистемическим горизонтом, делающим их будущее недетерминированным. В этой связи он вспомнил известную статью К. Поппера «Индетерминизм в квантовой физике и классической физике»<sup>1</sup>, в которой тот защищал тезис о принципиальной невозможности исключения индетерминизма и в то же время о возможности его объективистского объяснения. Аргумент Деннета о непредсказуемости личностного поведения таков: «Это происходит потому, что ни одна информационная система не может содержать в себе полную и истинную репрезентацию самой себя (выражена ли эта репрезентация в терминах физической установки или в какой-либо иной). Отсюда следует, что я не могу даже в принципе иметь все данные, исходя из которых я был бы в состоянии предсказать свое будущее (исходя из любой установки)... И именно в силу того, что я должен рассматривать себя как личность и как полноправную интенциональную систему, полной биографии моего будущего, которую по праву я должен был бы принять, не существует»<sup>2</sup>.

Тем не менее, замечает Деннет, сказанное еще не является доказательством невозможности полной деперсонализации в будущем. Если успехи механистического объяснения человеческого поведения сами по себе не лишают нас ответственности, они заставляют подходить к ней более прагматично. Еще сравнительно недавно единственным эффективным способом заставить людей делать что-либо в соответствии с желаниями других людей было обращение с ними как с личностями. На них воздействовали угрозами, ложной информацией, подкупам, однако все это были средства контроля и принуждения, эксплуатирующие рациональность. Установка дизайнера

или физическая установка

<sup>1</sup> Popper K. *Indeterminism in Quantum Physics and Classical Physics* // *British Journal of Philosophy of Science*, 1950.

<sup>2</sup> Dennett D. *Brainstorms*. P. 254—255.

278

не применялись потому, что казалось маловероятным получить на их основе практически значимые поведенческие результаты. Вторжение в жизнь современного человека таких вещей, как промывка мозгов, действующая на подсознание реклама (все они относятся к установке дизайна) и прямое физическое вмешательство в мозг с использованием препаратов и хирургических операций, впервые сделали по-настоящему актуальной задачу *выбора установки*. На этом уровне многие моральные вопросы получают разрешение, если подводятся под рубрику «рассмотрение личностей как менее чем личностей во имя их собственного блага»<sup>1</sup>.

А что если массовый гипноз отобьет у людей желание курить? Или заставит людей отказаться от убийств? А может страдающий человек в результате операции лоботомии превратиться в удовлетворенную собой личность? Одним словом, делает он вывод, вопрос о том, наносят или не наносят механические воздействия ущерб рациональности и подрывают ли они моральную ответственность, должен решаться конкретно и контекстуально. На самом деле в этой сфере больше вопросов, чем ответов.

Какие выводы можно сделать из содержания работы Деннета «Механицизм и ответственность»? Многие обсуждаемые им идеи высказывались самыми разными авторами, и его собственный вклад, по его же признанию, состоит в наведении теоретического порядка в существующем на сегодня эклектизме и выработке унифицированного основания для идей, касающихся отношений между ответственностью и механицизмом.

Достижение унификации, как заметил читатель, у Деннета связано прежде всего с применением его теории установок: физической, дизайна и интенциональной. Собственно говоря,

<sup>1</sup> Ibid. P. 255.

279

рассматриваемая нами статья «Механицизм и ответственность» представляет собой дальнейшее *развитие теории установок* с точки зрения их взаимоотношения в объяснении *поведения* систем — как личностных, так и не-личностных.

В более общем виде ее можно рассматривать как развитие *логического бихевиоризма*, выдвигаемого в качестве альтернативы традиционному ментализму и концепциям метафизически личностного. Важно не забывать, что сутью теории установок является разработка объективистской методологии (в его терминологии, «гетерофеноменологии») для деконструкции картезианского образа сознания и всех связанных с ним представлений — о центральном положении самости, метафизически личностном, свободе воли и ответственности.

Как мог заметить читатель, позиция Деннета в понимании ответственности все же *уклончива*. Оптимист, верящий в прогресс науки и философии, он выглядит не очень уверенным, когда от него требуется однозначный ответ на вопрос: «Возможна ли полная гегемония механицизма?» Слишком много вопросов остается без ответа. Нам представляется, его неуверенность связана с тем, что, изобретая инструменты для низложения картезианского онтологического дуализма, он постоянно наталкивается на воспроизведение других форм дуализма — гносеологического, методологического и др.

С одной стороны, он выступает и против усмотрения антагонизма между механицистским и целеполагающим (интенциональным) объяснениями, и против внешнего «присоединения» одного к другому. И в антагонизме, и в «присоединении» ему видится угроза единству знания — принципиально важному для него идеалу (и намерения опровергнуть «тезис Brentano»). Он определенно заявляет о *совместимости* и возможной плодотворной работе двух типов объяснения. Можно применять интенциональную установку по отношению к компьютерной машине, а механицистскую — к телесной организации людей. Во всяком

280

случае, сегодня у нас нет оснований говорить об исчерпанности интенциональной установки: в объяснении поведения систем лучшей стратегии не предложено. С другой стороны, Деннет не исключает правоту Куайна, что в будущем, «в какую-то следующую неделю», механицистское объяснение сможет вытеснить интенциональное и, соответственно, мы будем рассматривать личностей «как меньше чем личностей во имя их собственного блага». Иначе говоря, позицию Деннета, как и других радикальных материалистов, можно отнести к категории «обещающего материализма». Сторонники «обещающего материализма» настаивают на теоретической (не практической) возможности полного описания человека в физикалистских терминах (тела, поведения, мыслей), хотя и относят эту возможность в неопределенно далекое будущее.

Важно иметь в виду, что в рамках общей физикалистской парадигмы только первые две установки — установку дизайна и физическую установку — Деннет наделяет *онтологическим* статусом. Интенциональная установка не имеет его. В отличие от двух первых она *нормативна*, есть дело *решения и выбора, а не открытия*; она зависит от правил нашей социолингвистической практики. То же самое относится к личностной и моральной установкам. Для понимания их природы принципиально важно, что все они — нормативные *прагматические установки*. Критерием их принятия или непринятия является успех в предсказании и в контроле поведения без апелляции к тому, являются ли «онтологически реальными» имеющиеся у объекта верования и интенции. Что касается механицистских установок (физической и дизайна), то здесь дело обстоит по-другому. Реальность нейрофизиологических процессов (или электронов) нами открывается, а не выбирается и не изобретается. Иначе говоря, предложенные Деннетом установки по своему статусу не являются равноправными. Можно даже сказать, что картезианский онтологический (субстанциалистский) дуализм он заменяет дуализмом объяснительных установок.

281

Конечно, если последовательно придерживаться инструменталистской эпистемологии, между этими установками нет антагонизма и они не образуют дуализма. Если же придерживаться реализма, а Деннет претендует на это, то они есть. Когда он говорит о нереальности верований и реальности электронов, по сути он воспроизводит тот же самый антагонизм механицистского и целеполагающего объяснений, который он хотел преодолеть.

Различие в статусах установок по логике должно детерминировать движение Деннета в сторону признания гегемонии механицистского объяснения. Однако такой путь его не устраивает. Он приводит множество условий, с тем чтобы доказать не только нужность, но и неизбежность интенционального. На сегодняшний день механицистское объяснение нереалистично ни в отношении личностей, ни в отношении компьютеров. Оно и нежелательно, ибо ставит под вопрос рациональность нашей телесной организации; гегемония механицистского объяснения означала бы отказ рассматривать нас самих как существа, имеющие рациональный дизайн. В более поздних работах («Интенциональная установка») он пытается подвести под эти выводы более весомые аргументы и фактически трансформировал интенциональность из методологического принципа в онтологический: роль первичного дизайнера интенциональности он отдал Матери-Природе. Как мы уже говорили, у Деннета всегда присутствует двойная перспектива. С одной стороны, и интенциональность и целеполагание лишены онтологического статуса (это нормативные вещи, на онтологию может претендовать только механицистское объяснение). С другой стороны, он объективирует их, наделяя сами природные процессы телеологием и протоинтенциональностью.

Из приведенного выше текста видно, что Деннет, в отличие от многих других авторов, не отождествляет интенциональность с личностным и моральным. Тем более он не ищет оснований для нее ни в каузальных событиях, ни в «метафизически личностном». Интенциональное является предва-

282

рительным условием личностного и морального. Личностная или моральная установки надстраиваются над интенциональным и по своему значению уже понятия интенционального.

Различение более широкого понятия «интенционального» и более узкого понятия «морально личностного» нужно Деннету для того, чтобы, с одной стороны, ввести поведение механических систем, в частности компьютера, в сферу действия интенциональной установки, а с другой стороны, вывести его из сферы личностной установки, где действуют мораль и ответственность. Можно, конечно, вести умные беседы с роботом, тем не менее, советует он, все же не стоит включать его в наше человеческое сообщество и возлагать на него ответственность за плохие поступки. Если принять такое различение, то получается, что мы применяем два разных варианта рациональности: один — когда имеем дело с поведением компьютера, другой — когда имеем дело с поведением людей; первый будет рациональностью без ответственности, второй — рациональностью с ответственностью. Иначе говоря, мы постулируем дуализм рациональности. Если следовать физикализму Куайна, было бы логичнее распространить и на человека и на робота единые принципы рациональности и приписать роботу свойства личностного, моральности и ответственности. На такой радикальный шаг Деннет, разумеется, не идет. Однако он логически вытекает из его рассуждений. Одним словом, его борьба против картезианского дуализма с позиции физикалистского монизма имеет постоянное следствие: дуализм, устраненный в одном месте, возрождается в другом.

В заключение мы хотели бы сказать несколько слов о философской стилистике Деннета. Как и авторы, с которыми он полемизирует, он не склонен поднимать вопрос о механицизме и ответственности на «высокий» метафизический уровень и пользоваться пафосной морально-гуманистической риторикой. Его интересуют только доводы и контрдоводы, примеры и контрпримеры. Конечно, порою такая стилистика оставляет у читателя ощущение философской незавершенности. Вместе

283

с тем нельзя не признать ее практическую эффективность: в конкретных ситуациях выявляются теоретические изъяны, концептуальные загадки и неясности, возникающие при рассмотрении той или иной проблемы в разных, в том числе и виртуальных, контекстах. Такой метод является практически значимым для таких областей, как юриспруденция, которая нуждается не в риторике, а в обоснованных аргументах, позволяющих оценивать личность как вменяемую или невменяемую. Потребность в строгих доводах возрастает с расширением технических возможностей вторжения в мозг, манипуляции людьми, с клонированием организмов и т.д.



## Часть II. ДЕННЕТ И КРИТИКА

### I. Споры вокруг философии Деннета

Почти с уверенностью можно сказать, что Деннет сегодня занимает — наряду или сразу за Расселом и Витгенштейном — положение наиболее известного философа этого столетия. Хотя быть известным *за* свои идеи не то же самое, что эти идеи являются знаменитыми.

*Дон Росс*

Когда от меня потребуют провозгласить альянс с элиминативными материалистами, или с дуалистами свойств, или с реализмом, или с эмерджентизмом, или даже (да простит мне Бог) с эпифеноменализмом, или с чем-нибудь еще, я буду настойчиво сопротивляться таким требованиям.

*Дэниел Деннет*

Деннет не испытывает недостатка в критике. Масштабность и провокативность проекта деконструкции картезианского образа сознания и замена его «Моделью Множественных Наброс-

285

ков», метафизика интенциональности, построенная на идее природы как «Главном Дизайнере», концепция алгоритмов «Машины Дарвина», «Машины Тьюринга» и «Машины Сознания», активное участие в «Дарвиновских войнах», работа «на стыках» (или, как он говорит, «Living on the Edge»), не говоря уже о собственной склонности к полемике, делают его философию объектом разноречивых комментариев. В последние годы, благодаря доступности его работ широкой аудитории через Интернет, обсуждение его взглядов приобрело, можно сказать, «вселенский характер». Представители разных философских и научных направлений по-разному оценивают новации Деннета: у одних они вызывают одобрение, у других — недоумение и даже негодование.

Начнем с позитива. И у апологетов и у критиков удивление и восхищение вызывает информированность Деннета во всех областях, имеющих касательство к исследованию интеллекта, его подчас профессиональное знание ситуации в когнитивных дисциплинах, науках об искусственном интеллекте, биологии и др. («Чистые» философы признаются, что им трудно спорить с ним, когда тот ссылается на результаты экспериментов, о которых могут судить только профессионалы.) Все отмечают оригинальность философского стиля Деннета, сочетающего «инженерный подход» с богатым философским воображением, его умение в полемике приводить примеры и контрпримеры, видеть в привычных вещах философские загадки и вовлекать читателя в интересное приключение по их разгадыванию.

Практически все пишут о смелости его проекта — представить непротиворечивую, на самом современном уровне материалистическую концепцию сознания. Приведем мнение одного из оппонентов Деннета — профессора Оксфордского университета Джона Фостера, к слову сказать, убежденного дуалиста. Хотя Деннету не удалось положить Декарта на лопатки, говорит он, «в действительности может случиться так, что деннетовская догматическая расправа с дуализмом, обрамленная его красочной экспозицией функционалистской альтер-

286

нативы, окажется лучшим из всего того, что может сделать материалист»<sup>1</sup>.

Высоко оценивает его творчество Ричард Рорти. Он называет Деннета глашатаем, наряду с Витгенштейном, Райлом и Куайном, «копернианского поворота» в философии сознания. Используя для деконструкции главного персонажа европейской культуры — самости — компьютерную аналогию, в главном он продолжил прагматическую и бихевиористскую линию этих философов на замену трансцендентного Я субъектом говорящим и прагматически действующим. Хотя Деннет и противится упоминаниям его имени рядом с постмодернистами, по большому счету, считает Рорти, он идет в том же направлении, что и европейские деконструктивисты, освобождающие наш язык от модернистских псевдореалий<sup>2</sup>.

С 1993 по 2002 год вышли четыре сборника, посвященные обсуждению взглядов Деннета: «Деннет и его критики. Демистифицируемое сознание», 1993; Материалы симпозиума по книге

«Сознание объясненное», 1993; «Философия Деннета: общая оценка», 2000; «Дэниел Деннет», 2002<sup>3</sup>. Среди приняв-

<sup>1</sup> Foster J. Dennett's Rejection of Dualism. Symposium // Inquiry. V. 36. № 1/2. 1993. P. 30. Книга Фостера «Нематериальная самость: защита картезианской дуалистической концепции сознания» (Foster J. The Immaterial Self.: a Defence of the Cartesian Dualist Conception of Mind. L., 1991) представляет собой, как он говорит, «ответ на волну физикализма».

<sup>2</sup> Более подробно о сходстве и различиях позиций Деннета и Рорти мы будем говорить в разделе «Деннет versus Рорти...» См. также: Деннет Д. Почему нам важно понимать это правильно. Постмодернизм и истина // Настоящее издание. Приложение 1. С. 477—492; Юлина Н.С. Д. Деннет о вирусе постмодернизма // Вопросы философии. № 8. М., 2001. С. 93—100.

<sup>3</sup> Dennett and His Critics. Demystifying Mind. Ed. by Bo Dahlbom. Cambridge MA, 1993; Daniel Dennett's «Consciousness Explained». Symposium // Inquiry. V. 36. № 1/2. 1993; Dennett's Philosophy: A Comprehensive Assessment. Ed. by D. Ross, A. Brook, and D. Thompson. Cambridge MA, 2000; Daniel Dennett. Ed. by A. Brook and D. Ross. Cambridge MA, 2002.

287

ших в них участие авторов известные философы, зоологи, психологи, специалисты по искусственному интеллекту, нейронауке, психофизике, когнитивной этологии и др. При обозрении их комментариев (все они написаны в проблемном ключе) бросается в глаза прежде всего живой интерес ученых к дискуссиям в философии сознания и их хорошая философская подготовка. Второе — это многообразие перспектив, в которых авторам видится философия Деннета.

По мнению Бо Дальбома, редактора первого сборника «Деннет и его критики», специфика философии Деннета такова, что по отношению к ней можно использовать множество «измов»: радикализм, функционализм, номинализм, конструктивизм, бихевиоризм, верификационизм, холизм, сциентизм, реляционизм, интерпретивизм и др. Однако все «измы» стягиваются в один узел — бихевиоризм, связывающий Деннета с Витгенштейном, Райлом и Куайном, с одной стороны, и Ницше, Рорти и Деррида, с другой<sup>1</sup>. Другая оценка у Джона Хогелэнда. Будучи специалистом по Хайдеггеру, он увидел в философии Деннета грандиозный онтологический проект — теорию бытия, по масштабам и амбициям напоминающий проект немецкого философа. Введение им в концепцию интенциональности идеи «Реальных Паттернов» есть не что иное, как возведение онтологической конструкции, нацеленной на выявление бытийственных черт мира. Главное в понятии «интенциональная установка» не столько экстернализм, сколько способ показа единства бытия<sup>2</sup>. По мнению Колин МакГинн, концепцию Деннета лучше всего отнести к «инструментальному логицизму»: сознание у него представлено в виде логической когнитивной машины, а интенциональная установка в его руках не что иное, как применение логики и ее норм для объяснения производства мысли и типов поведения на основе посылки рациональности. Прида-

<sup>1</sup> Dennett and His Critics. P. 6.

<sup>2</sup> Ibid. P. 53.

288

ние рациональности конституитивной и нормативной роли, считает он, сближает его с другими ментальными логицистами, в частности с Дональдом Дэвидсоном<sup>1</sup>. В чем-то сходную оценку дает Пол Черчленд. Он относит Деннета к «скрытым гегельянцам», внимание которых приковано к динамике концептуальной истории, переплетающейся с динамикой природных процессов<sup>2</sup>. Джерри Фодор увидел в философии Деннета «амбициозный метафизический проект». Его эпистемология, в особенности интерпретивизм интенциональной установки, считает он, сконструирована в традиции инструментализма, пороки которого он безуспешно пытается устранить с помощью эволюционизма<sup>3</sup>. Самым язвительным образом о философии Деннета отозвался Томас Нагель: это, говорит он, Джилльберт Райл, скрещенный с журналом «Scientific American»<sup>4</sup>.

Весьма любопытна оценка «социологиста» Бо Дальбома: в концепции сознания Деннета ему видится «замечательный способ суммировать идею анархического либерализма и идею свободного рынка! ...Отрывая нас от метафоры Картезианского Театра и завлекая метафорой

Модели Множественных набросков, он явно идет в ногу с более общим идеологическим сдвигом, имевшим место в 80-х годах как в мышлении менеджмента, так и, в определенной степени, его практике. Деннет приглашает нас применить «постфордизм» к сознанию, прекратить представлять сознание в виде централизованной бюрократической организации типа производственных линий на заводах Форда и начать думать о сознании скорее как децентрализованной, органической, гибкой, организации»<sup>5</sup>.

<sup>1</sup> Ibid. P. 87.

<sup>2</sup> Dennett Daniel. Ed. by A. Brook and D. Ross. Cambridge MA, 2002. P. 64.

<sup>3</sup> Dennett and His Critics. P. 77.

<sup>4</sup> Dennett D. Review of Nagel «Other Minds». Internet: <http://ase.tufts.edu/cogstud/pubpage.htm>. P. 2.

<sup>5</sup> Dahlbom Bo. Mind is Artificial // Dennett and His Critics. Demistifying Mind. P. 164.

289

Не менее разноречивы оценки философии Деннета со стороны ученых. Одни упрекают его за незнание каких-то новых научных фактов, другие — за интерпретацию научных фактов через призму философских гипотез. Критикуют и за излишний социологизм, и за крен к биологицизму. Следует отметить, что многим ученым импонирует атака Деннета на феноменализм. Известные психологи Т. Марсель и Б. Баарс, нейрочеловек М. Кингсборн признают, что его критические аргументы против существования квалиа оказались методологически полезными в конкретных исследованиях. В качестве позитива учеными чаще всего выделяется прикладная значимость концепции «интенциональных установок». По словам известных психологов Р. Гриффина и С. Барон-Козна, применение методологии интенциональных установок дало хорошие результаты в объяснении психологического развития детей и некоторых психических патологий. Когнитивные этологи-экспериментаторы Р. Сейфарт и Д. Чини применили интенциональную установку при объяснении поведения приматов и нашли ее весьма эффективной. Специалист по экономике и когнитивным наукам Д. Росс считает, что методология установок перспективна в социальных науках при объяснении социального поведения людей и институтов, где постоянно возникает необходимость соотнесения их действий с интенциями или резонами, а не только с материальными причинами. Применима она и в экономической сфере, в теории игр и ряде других областей, где важно учитывать намерения людей<sup>1</sup>. И конечно, все они видят большой позитив в том, что авторитетный американский философ работает, повернувшись лицом к науке и рефлексии о науке. Очень точно об этом сказали Р. Гриффин и С. Барон-Козн: «Деннет не посылает

<sup>1</sup> Griffin R. and Baron-Cohen S. The Intentional Stance: Developmental and Neurocognitive Perspectives // Daniel Dennett. P. 83—116. Там же: Seyfart R. and Cheney D. Dennett's Contribution to Research on Animal Mind. P. 117—139; Ross D. Dennettian Behavioral Explanations and the Roles of the Social Sciences. P. 140—183.

290

путаные послания, он просто пособляет хорошей науке. Деннет — поборник науки, хотя ему хорошо известны многие потенциальные минные поля, которые философия взяла под свою ответственность, такие, например, как проблема сведения значения к синтаксису мозга, особо острая для тех из нас, кто свободно рассуждает о репрезентациях. И все же, вместо заявлений о непроходимости минного поля, что характерно для многих работающих в области философии, Деннет протягивает нам схемы, показывающие как можно пройти этот потенциально опасный курс»<sup>1</sup>.

Разноречия в оценках творчества Деннета, как нам кажется, во многом объясняются стилистикой его философствования. Она отличается от принятой в англоязычной традиции манеры рассуждения с ясным изложением тезисов, приведением четких доказательств, обоснованием принятых критериев и т.п. Например, среди его многочисленных работ нет ни одной с изложением принципов его метафизики (хотя он и не отрицает, что таковая у него есть); комментаторам приходится, проделывая трудоемкую работу, «вычитывать» ее, что не может не сказываться на характере интерпретаций.

Разумеется, Деннету трудно согласиться с однозначными идентификациями его позиции путем подведения ее под какой-либо «изм». Он все время поясняет, что его интерес к деталям и примерам, объяснение которых не укладывается ни в одну известную теорию, не позволяет легко и просто ее каталогизировать. (Действительно, девизом его работ можно назвать «Дьявол кроется в деталях».) Конечно, говорит он, это осложняет дело моим интерпретаторам, и посему они изображают его позицию такой, какой она им кажется, а не такой, какой она является.

Особенно часто Деннет слышит упреки в том, что он не определился в своих онтологических воззрениях и не изложил их в систематической форме. Этого я не буду делать,

<sup>1</sup> Griffin R. and Baron-Cohen S. *The Intentional Stance: Developmental and Neurocognitive Perspectives*. Op. cit. P. 107.

291

отвечает он. Не из-за негативного отношения к онтологии; работа не только философии, но и науки отнюдь не является онтологически нейтральной, чаще всего это просто нерелексивно принятая онтология, но она оказывает свое влияние на практику. Например, одной из причин отсутствия прогресса в когнитивных науках является царящая там онтологическая неразбериха. Выход состоит не в том, чтобы приостановить работу науки до прояснения онтологии, а, скорее, поступать так, как это предлагал Отто Нейрат — ремонтировать лодку онтологии, находясь в плавании, то есть замечая в ней дыры, ведя поиск аналогий, приводя подтверждающие или опровергающие примеры и т.д. Любителю метафизики и онтологии он отвечает так: «Смотри сюда, и когда тебе станут ясными такие не-мистические случаи, как улыбка, возможности, центры гравитации и голоса, тогда (и только тогда) я сочту себя обязанным отвечать на твои вопросы. А до того онтологическая деятельность не представляется мне чем-то, во что я должен втягиваться»<sup>1</sup>. Более интересны по-настоящему загадочные вопросы о природе сознания, самости, свободе воли, ответственности.

Разумеется, самая главная причина разноречивых оценок творчества Деннета кроется в существовании большого разброса позиций по ключевой проблеме «Что такое сознание и как его можно исследовать?» Нет консенсуса ни по одному из связанных с ней вопросов. Существует ли сознание как особая реальность? Является ли оно эпифеноменальной «добавкой» к физической Вселенной или Вселенная монистична, непрерывна, и все что мы знаем, или можем узнать, постигается обычными, принятыми в науке, рациональными способами? Если это все же «добавка», то как она возникла: как инновация биологической эволюции (ее можно называть как угодно — «эмерджентией», «супервенцией» и др.), или она — чистое творение социальности, языка и культуры. Можно ли сказать, что сознание есть особое «свойство», или оно — это только функ-

<sup>1</sup> Dennett and His Critics. P. 212.

292

циональные отношения, которые не сводятся ни к каким материальным или идеальным свойствам? Причем при любом толковании альтернатив загадкой остается механизм скачка от довербального уровня мозговых процессов к вербальному уровню сознательных процессов, от синтаксиса к семантике.

Конечно, в мире существует множество тайн — Вселенной, материи, Природы и др. Но в отношении этих тайн по меньшей мере есть согласие, что их исследование следует проводить рациональными, принятыми в науке способами. Этого нельзя сказать про тайну сознания. За полвека интенсивных дискуссий в англоязычной философии сознания в ней не сложился консенсус в отношении методологии исследования сознания: подходить ли к нему «изнутри» (интернализм) или «извне» (экстернализм), «снизу» (из биологии) или «сверху» (из языка и культуры). Или вообще отказаться от всяческих попыток и оставить разгадку тайны сознания следующим поколениям. Некоторые философы, разочарованные неудачами, говорят о кризисе науки и даже рациональности, коли с их помощью невозможно объяснить сознание. Другие связывают прояснение загадки сознания с революцией в физике, в частности в квантовой механике. Сегодня у нее не меньше загадок, чем в философии сознания, однако по крайней мере

известно, на каком пути искать их разгадки. Есть и противоположные мнения. Некоторые авторы полагают, что если квантовая механика сегодня не в состоянии пролить свет на тайну сознания, то, возможно, новый философский взгляд на сознание поможет разобраться с трудностями квантовой механики. Во всяком случае, распространено убеждение, что между квантовой механикой и философией сознания есть связь уже потому, что многие трудности и там и здесь упираются в понимание опыта и наблюдателя (субъекта).

Можно сказать, что отсутствие консенсуса в философии сознания по большому счету отражает противоречивые позиции в отношении объективистских методов. Практически все комментаторы — не только дуалисты, но и материалисты-монисты, — признают наличие больших трудностей у

мате-

293

риализма с его объективистскими методами при объяснении феноменального, субъективного, самостно-личностного опыта. Несмотря на накопление в когнитивных науках массива данных, свидетельствующих о ложности интроспекции, о том, что люди обманываются в отчетах о своих ментальных состояниях и мистифицируют роль своего *Я*, все же человеку трудно сомневаться в реальности его ощущений и чувств. Признание их только кажимостями чересчур контринтуитивно. Поэтому вновь и вновь возникают сомнения относительно возможности объективного описания субъективности. Дж. Сёрль, Т. Нагель, Дж. Фодор и многие другие предупреждают, что говорить о девальвации нашего интуитивного представления о самом себе (нашей фолк-психологии), учитывая ограниченность современных объективистских методов, преждевременно. Дэвид Чэлмерс, обзревая более чем полувековую дискуссию о сознании, в которой проиграно множество гипотез, считает очевидным, что без эпифеноменального не обойтись и, подавив в себе интенцию к монизму, легализовать «дуализм свойств» или «дуалистический натурализм»<sup>1</sup>. Введенный Чэлмерсом термин «дуалистический натурализм» звучит, конечно, необычно, тем не менее в чем-то он точно маркирует смысл многих позиций, настаивающих на том, что сознание представляет собой особое, нередуцируемое *свойство* мира, а не только функциональные отношения свойств.

Одним из критических вопросов в дебатах вокруг деннетовского проекта объяснения сознания без обращения к ментальному является эпистемологический вопрос: что значит *объективно* описывать процессы сознания? На этот счет существуют по меньшей мере две основные позиции, которые можно назвать панобъективизмом и умеренным объективизмом (с поправками на субъективность).

Напомним, что спор о преимуществах методологического монизма (физикалистский панобъективизм) и методологичес-

<sup>1</sup> Chalmers D. The Conscious Mind. In Search of a Fundamental Theory. Oxford, N.Y., 1996. P. 123—212.

294

кого дуализма (объективизм с ограничениями в отношении субъективности) имеет давнюю историю. В англоязычной философии глашатаем панобъективизма был, конечно, Куайн. Однако не все физикалисты были столь уверены в непререкаемости физикалистской панобъективистской парадигмы в объяснении субъективности и личностного. Например, У. Селларс говорил о «трагизме научного образа мира», не могущего в своих объяснительных схемах охватить личностное отношение человека к миру. Были и другие сциентистски и натуралистически мыслящие философы, высказывавшиеся за умеренный объективизм.

Как бы то ни было, можно констатировать, что к началу XXI века в англоязычной философии сознания произошел раскол, одной из причин которого является асимметрия между нашим прямым доступом к сознанию и внешним, объективным, теоретическим его описанием. Деннет образно говорит об образовании на поле философии «двух команд». Первую команду образовали «перволичностники», настаивающие на реальности сознания. Это Томас Нагель, Джон Сёрль, Джерри Фодор, С. Харнад, и лидирует в ней на сегодня Дэвид Чэлмерс. В команду «третьеличностников» и объективистов, отрицающих такое бытие, он включил Уиларда Куайна,



Ричарда Рорти, Дугласа Хофстадтера, Пола и Патрицию Черчленд, Энди Кларка, Дэвида Розенталя, Джилберта Хармана, Стивена Стича и, конечно, себя как ее лидера.

Поскольку концепцию Деннета трудно представить без непрерывных дискуссий с коллегами, нужно, строго говоря, нарисовать панорамную картину, которая именовалась бы приблизительно так: «Деннет versus Скиннер, versus Рорти, versus Нагель, versus Сёрль, versus Фодор, versus Дэвидсон, versus Ким, versus Патнэм, versus Фостер, versus Локвуд, versus МакГинн, versus Чэлмерс и др.». Создание такой картины выходит за рамки задач нашей работы. Для того чтобы читатель в какой-то мере получил представление о некоторых содержательных узлах ведущейся дискуссии, мы включили в материал глав ряд критических комментариев оппонентов (Дж. Фодора, Э. Клар-

295

ка, П. Черчленда, Дж. Сёрля и др.). Ниже мы более подробно остановимся на полемике Деннета с Нагелем, Сёрлем и Рорти. В заключение сопоставим взгляды на сознание Деннета и Поппера. Выбор этих философов обусловлен тем, что они сосредоточили критику на болевых точках концепции Деннета: понимании сознания, объективности, компьютерных аналогий, реализма и др. Имеет смысл начать с критических аргументов Нагеля, поскольку они сконцентрированы на проблеме объективное — субъективное и более явно выходят на общемировоззренческий уровень.

## II. НАГЕЛЬ VERSUS ДЕННЕТ: «ОБЪЕКТИВНАЯ ФЕНОМЕНОЛОГИЯ» ПРОТИВ ПАНОБЪЕКТИВИЗМА

Амбиция трансцендентности Нагеля — это не безграничная преданность интеллектуальной честности, как он думает, скорее это мягкосердечная тоска по невозможной стабильности и порядку — разновидность тоски, подноготную которой Уильям Джемс вскрыл в первой главе своего «Прагматизма».

*Ричард Рорти*

Критическое отношение Томаса Нагеля к взглядам Деннета проявилось еще в начале 70-х годов, когда он опубликовал рецензию на его первую книгу «Содержание и сознание», выразив сомнение относительно возможности девальвации феноменологии. Позднее он назвал «уловкой» его тезис, что констатацию о наличии сознания у разумно ведущего себя взрослого человеческого существа следует считать только *верованием*; на самом деле, писал он, такое утверждение есть *вывод из очевидности*, а не только *вера в очевидность*. В первом случае сознание пропадает, во втором оно постулируется. В 1974 году Нагель опубликовал статью «Что это такое, быть летучей мышью?», в которой атрибутивным качеством живого существа назвал бытийственное отношение к миру. Его собственный ответ на поставленный вопрос свелся к тому, что нам не дано представить, *что же это такое* — *быть* мышью (или дельфином, или человеком, или кем-либо еще)? Представить объективистски, что может

297

означать опыт субъекта, абстрагируясь от его специфической точки зрения, невозможно. «Любой сдвиг к большей объективности — то есть к меньшей зависимости от специфической точки зрения — не приближает нас к реальной природе феномена: он отодвигает нас от него»<sup>1</sup>. Сравнительно небольшая статья Нагеля, прозвучавшая диссонансом в оптимистическом хоре материалистических теорий, бросила вызов устойчивому мнению о всевластии каузальных интересубъективных объяснений. Она получила широкий отклик в литературе; многие не только философы, но и ученые сочли приведенные в ней аргументы убедительными.

В книге «Взгляд из ниоткуда» (1986) Нагель продолжил эту тему. Существенным изъяном объективистских подходов к опыту, говорит он, является даже не их интересубъективность *per se*, а их уход от вопроса, что значит быть самим собой. Панообъективизм в его физикалистском или функционалистском варианте принципиально не в состоянии справиться с этим вопросом. В объективизме «сделана попытка посмотреть на мир не из какого-то места внутри него или с преимущественной точки особого типа жизни и осознания, но именно из ниоткуда, из никакой специфической формы жизни»<sup>2</sup>. «Взгляд из ниоткуда» оставляет за бортом очень многое, что не

может быть адекватно понято с максимально объективной позиции, как бы далеко она ни расширяла свои границы.

Главная эпистемологическая посылка его контраргумента состоит в следующем: ментальные феномены не могут быть объяснены просто потому, *что они входят в само объяснение*. Приписывая эту способность личности, мы интерпретируем ее деятельность в терминах собственной способности. В отличие, скажем, от феномена света, который может быть представлен и объективно (внезапная разрядка электрической

<sup>1</sup> Nagel Th. What is it to be a Bat? // Nagel Th. Mortal Questions. Cambridge MA, 1979. P. 174.

<sup>2</sup> Nagel Th. The View from Nowhere. N.Y., Oxford, 1986. P. 7—8.

298

энергии) и субъективно (восприятие его нами как ослепительной вспышки), опыт сознания не допускает никакого различия между тем, как событие представлено сознанию, и каково это событие есть само по себе. Опыт сознания существует как определенное явление и всегда как явление некоторому субъекту. «Поскольку я не могу выпрыгнуть из нее (своей перспективы — *Н. Ю.*) вовне, как я могу надеяться выпрыгнуть вовне и объяснить соответствующие вещи у кого-либо другого»<sup>1</sup>. Оптимисты, говорит он, обманывают себя, надеясь объективистски описать единство воспринимаемого и воспринимающего. Объективистская слепота негативно сказывается на современной философии сознания, в которой получает распространение то одна, то другая экстерналистская теория ментального — от физикализма до функционализма. Экстерналисты считают образцом реальности мир, описываемый физикой, — наукой, наиболее отчужденной от специфически человеческого видения мира. Упование на дарвиновский эволюционизм тоже самообман. Торжество «дарвиновского империализма», наблюдаемое в англоязычной философии, Нагель оценивает как результат поиска емких и всеохватывающих объяснительных моделей при боязни впасть в религию. Хотя и в дарвинизме, и в современной физике на самом деле таится религиозная угроза, большинству атеистов она кажется менее пугающей. Нагель категорически не согласен с Деннетом, что эволюционный аргумент может служить оправданием и обоснованием рациональности (и веры в истину); эволюционный натурализм, замечает он, «всегда казался мне до смешного неадекватным»<sup>2</sup>, не говоря уже о том, что эволюционизм, как и физикализм, чреват телеологием и фатализмом. Получается, что ментальные способности, которые потребовались нашим предкам для создания каменного топора, автоматически включали в себя способности к реализации последующих шагов, приведших в конце концов к созданию водородной

<sup>1</sup> Ibid. P. 76.

<sup>2</sup> Nagel Th. The Last Word. N.Y., Oxford, 1997. P. 75.

299

бомбы. Или, как в фильме «2001» Стэнли Кубрика, в древнем орудии из кости каким-то неведомым образом уже существовала тенденция его трансформации в космический корабль. Одним словом, эволюционное объяснение нельзя рассматривать как независимое от разума оправдание разума; оно не часть «первой философии», а только часть нынешнего изменчивого, не развернувшего свои потенциалы научного мировоззрения.

Есть ли альтернатива физикалистскому и дарвинистскому «империализму»? Нагель полагает, что есть. Эту альтернативу он называет «объективистской феноменологией». Он не дает ей конкретной расшифровки (и не поясняет детали стыковки с наукой). В ней просто декларируется основная посылка: интеллектуальная честность требует признания, что некоторые вещи не могут быть объяснены, т.к. они входят в само объяснение. Философу-объективисту, оговаривает он, нет необходимости беспокоиться по поводу несовместимости феномена сознания с законами термодинамики: достаточно признать, что «субъективность сознания является нередуцируемой чертой реальности, без нее мы не можем создавать физику или что-либо еще, и посему она должна занимать столь же фундаментальное место в любом уважаемом воззрении на мир, какое занимают материя, энергия, пространство, время и числа»<sup>1</sup>. Общий его вывод сводится к следующему: «Вполне возможно принять мировоззрение, которое не объясняет все на свете в терминах квантовых полей, но и не содержит необходимую веру в Бога. Если естественный порядок может

включать универсальные, математически красивые законы фундаментальной физики того типа, который нам известен, почему же он не может включать такие же фундаментальные законы и ограничения, о которых мы ничего не знаем, но которые, возможно, совместимы с законами физики и делают постигаемым развитие организмов, наделенных сознанием, некоторые из которых способны раскрыть — в ходе длительных коллек-

<sup>1</sup> Ibid. P. 7—8.

300

тивных попыток — ряд фундаментальных истин относительно этого самого естественного порядка?»<sup>1</sup> А пока нам это неизвестно, следует ограничиться констатацией, что «в настоящее время у нас нет концептуальных средств, которые позволили бы нам понять, каким образом субъективные и физические свойства могут быть одновременно существенными сторонами единой сущности или процесса»<sup>2</sup>, то есть человека.

Из приведенных рассуждений Нагеля читатель может судить, что, несмотря на негативизм в отношении нынешних объективистских объяснений, альтернатива Нагеля достаточно строга: в ней не фигурируют картезианский «дух», трансцендентальное эго, интроспекция, вчувствование и т.п. По большому счету, его позицию тоже можно отнести к физикализму и объективизму. В отличие от Деннета, утверждающего, что мы должны объяснить сознание сегодня с помощью имеющихся в нашем распоряжении средств, достижение подлинного объективизма у него переносится в неопределенно далекое будущее.

На протяжении многих лет и в разной связи Деннет неоднократно возвращался к исходным идеям статьи Нагеля. Она служила ему как бы оселком для оттачивания аргументов против феноменализма и субъективизма. (В отличие от Нагеля, он предпочитает не входить в общие рассуждения относительно метафизических достоинств физикалистской или эволюционистской парадигм. В характерной для него манере он обращается к деталям и примерам.) В представленных нами выше главах мы уже приводили ряд доводов Деннета, в данном разделе мы суммируем главные.

Нагель, говорит он, декларирует «невыразимость внутреннего» в манере «папской непогрешимости», не называя ни одной черты субъективности мыши, которая либо сегодня, либо со временем не может стать объективно «выразимой». Вопрос

<sup>1</sup> Ibid. P. 131—132.

<sup>2</sup> Нагель Т. Мыслимость невозможного и проблема духа и тела // Вопросы философии. 2001. № 8. С. 105.

301

«Что такое быть летучей мышью?» на самом деле имеет негативный ответ: «Ничего». Если, как утверждают физиологи, конституция и функциональная организация мыши биологически эффективны, тогда иметь внутреннюю субъективность (свой взгляд на мир) ей просто ни к чему. То же самое можно сказать о человеке. В связи с этим можно вспомнить историю удивительного успеха производства по обогащению урана в Оук Ридж, организованного в рамках Манхэттенского проекта по созданию атомной бомбы. Тысячи людей ежечасно были заняты в контроле сложнейших процессов, делая все в срок и внося усовершенствования с учетом огромного количества составляющих условий. А между тем у них не было ни малейшего подозрения на то, что же они делали. Они не только не догадывались, что вносят вклад в создание атомной бомбы, но даже не знали, обогащают ли они уран или делают что-то другое. Можно сказать, что их девизом было: «Не спрашивайте меня, я только работаю здесь». То же самое можно сказать о мышце или мозге: не спрашивайте, где в них там прячется субъективность, смотрите только, как они работают. О субъективности мыши можно было бы говорить только тогда, когда свой взгляд на мир она выразила бы вербально, обмениваясь с помощью слов своим опытом с опытом других мышей.

Называя субъективность неотчуждаемым атрибутом живого существа (мышцы или человека), Нагель просто повторяет философов прошлого, рассуждавших о «внутренних сущностях» вещей, не постигаемых через явления. Мой же слоган, говорит Деннет, таков: «Если вы объяснили все,

что имело место, вы объяснили все»<sup>1</sup>. Собранная информация о всех возможных связях и отношениях объекта достаточна для его презентации и не нуждается в «добавке» в виде сущности или «самобытия», ускользающего от обычного рационального исследования.

Что мы ждем получить от объяснения сознания, и вообще, что мы можем ожидать от объяснения чего-либо? Эпистемоло-

<sup>1</sup> Dennett and His Critics. 1993. P. 210.

302

гически правильным ответом на такой вопрос, считает Деннет, является следующий: в любом удовлетворительном объяснении explanandum не может быть частью explanans. В философии сознания методологически нежелательно постулировать ментальное при объяснении ментального; в этом случае мы не выпрыгнем из замкнутого круга и обречены вечно топтаться вокруг тайны. Наука никогда не объяснила бы магнетизм, если бы в свои гипотезы изначально вводила бытовавшее когда-то понятие «аттрактант». Все явления, в том числе процессы сознания, должны объясняться на основе внешне фиксируемых, интересубъективно удостоверяемых факторов, независимо от их восприятия. Иначе говоря, надо действовать так, как действуют ученые. Дать объективную или научную характеристику какому-либо событию, значит, выйти за кажимость, абстрагироваться от его восприятия, от индивидуальных особенностей наблюдателя, от ценностных предпочтений и искать объективные каузальные связи «с позиции третьего лица».

Наблюдающаяся у философов сопротивляемость объективизму, говорит Деннет, очень часто коренится не в их теоретических убеждениях, а в глубинном предрассудке о неподвластности тайны сознания человеческому разуму, в особом психологическом настрое, когда, подойдя к самой критической точке, они в нерешительности останавливаются и говорят: «А здесь имеет место вечная загадка», и наука либо должна ждать своего повзреления либо ретироваться. Нагель, говорит Деннет, очень хорошо пишет о разуме и рациональности, и с его словами нельзя не согласиться. Однако его апелляции к «феноменологии субъективности» есть отступление от оправдавших себя принципов рациональности и от свойственного научной культуре оптимизма. На самом деле нет смысла в сомнении относительно хорошо работающих в науке стандартных принципов, каким бы релятивистским истолкованием они ни подвергались и какие бы ограничения они сегодня ни проявляли в толковании субъективного опыта. Стандарты, которым следует журнал «Scientific American», оправдывают себя в жизни.

303

В 1996 году Деннет опубликовал рецензию на книгу Нагеля «Other Minds». В целом он позитивно оценил его рефлексии по поводу «других умов», творивших философию в XX веке. Что касается «умов» в философии сознания (в том числе Деннета), то здесь он оказался не на высоте. Объяснение этому Деннет видит в том, что проблемы сознания он всегда воспринимал в этической перспективе. «Интерес Нагеля к сознанию всегда в большей мере подогревался его интересом к роли сознания в этике, нежели импульсами, идущими от нейрофизиологии. Он является моральным психологом, а не когнитивным ученым»<sup>1</sup>. Не имея вкуса к науке, он предпочел работать в традиционном каноне<sup>2</sup>.

Заклячая полемику Нагеля и Деннета, хотелось бы обратить внимание на последнее высказывание. Ибо суть полемики не только в убедительности или неубедительности сталкивающихся аргументов. Корни расхождения взглядов этих мыслителей глубже: в понимании, чем должна, а чем не должна заниматься философия, и вообще чего мы ждем от нее — объективного знания о мире или этической помощи во взаимодействии с миром? (В разделе «Деннет versus Рорти» мы покажем, что решение этого же метафилософского вопроса развело по разные стороны баррикад значительно более близких по умонастроению авторов.)

<sup>1</sup> Dennett D. Review of Nagel «Other Minds». Internet: <http://ase.tufts.edu/cogstud/pubpage.htm>. P. 1.

<sup>2</sup> Традиционализм Нагеля Деннет объясняет характером идейных влияний в период получения им философского образования. После окончания университета Корнуэлл Нагель готовил диссертацию в Оксфорде под руководством Джона Остина. В отличие от меня, говорит Деннет, получившего «куайновскую» натуралистическую закалку в Гарварде, которую не смогли

разрушить ни Райл, ни Оксфорд, Нагель воспринял характерную для Оксфорда того времени традицию отстраненности от науки, хотя и не пал жертвой другой традиции — лингвистической. Отстраненность от науки во многом определила его представление о философии как о дисциплине, развивающейся благодаря своим собственным ресурсам. (Ibid. P. 3).

304

### III. Сёрль versus Деннет: «Аргумент Китайской Комнаты»

Как исследователю AI, мне, возможно, трудно уловить различие между взглядами Деннета и Сёрля... Ведь оба согласны, что мозг является машиной... Что касается синтаксиса и семантики программ, то они выражены в очень похожих терминах, и никому не удастся построить философскую теорию о функционировании мозга или сознания на различии между ними.

*Йорик Уилкс*

#### I. Материализм и «онтология субъективного»

В 1992 году с небольшим временным отрывом от публикации книги Деннета «Сознание объясненное» (1991), доказывающей, что наше обычное представление о сознании есть сплошная иллюзия, вышла в свет книга Джона Сёрля «Открывая сознание заново»<sup>1</sup>, обосновывающая альтернативный тезис — мы должны доверять нашему здравому смыслу и верить в существование сознания. Деннет откликнулся на нее

<sup>1</sup> См. рецензию на нее: Юлина Н.С. Открывая сознание заново // Вопросы философии, 2003. №3. С. 186—191.

305

рецензией<sup>1</sup>. В 1998 году Сёрль опубликовал книгу «Тайна сознания»<sup>2</sup> с опровержением аргументов Деннета. Разумеется, Деннет тоже не остался в долгу.

Выступление двух выдающихся философов США, рисующих образ сознания с существенно разных, если не диаметрально противоположных позиций, драматизировало ситуацию и в философии сознания, и в когнитивной психологии. В связи с этим небезынтересно обратиться к продолжению их полемики. Мы не будем касаться всех затронутых в ней тонкостей и обратимся к общим вопросам. В первом разделе мы остановимся на различии их позиций в толковании материализма (и методологий объективизма и субъективизма), во втором — на философском толковании искусственного и естественного интеллекта.

В отличие от Нагеля, Сёрль чутко следит за тем, что происходит в науке, в нейрофизиологии, теориях искусственного интеллекта, и сам является практикующим когнитивным ученым. Тем не менее материализм ему видится по-иному, нежели Деннету. Сёрль не приемлет ту особую традицию материализма, которая сложилась в англоязычной философии сознания за последние полвека под влиянием бихевиоризма, теории тождества, элиминативизма и функционализма. Он с сожалением констатирует, что последний (в разных его вариантах) стал в ней доминирующей силой, своего рода ортодоксией, главный принцип которой — замена ментального на функциональное. Центральная доктрина функционализма состоит в том, что сознание представляет собой информационно-процессуальный (функциональный) уровень, опосредующий как феноменологический (личностный, субъективный), так и нейрофизиологический уровни. Главная цель доктрины — выведение за скобки субъекта, наделенного ментальностью. Сложилась разные формы функционализма, но все они в равной степени оказывают дезориентирующее влияние на исследование сознания.

<sup>1</sup> Journal of Philosophy. V. XC. №4. April 1993. P. 193—205.

<sup>2</sup> Searle J.R. Mystery of Consciousness. N.Y., 1998.

306

ния. Поэтому требуется детальное их опровержение. Сёрль оговаривает, что опровержение вовсе не означает, что он бросается в объятия дуализму или гуссерлианству. Он просто призывает произвести ревизию тому, что сделано в философии сознания, с тем чтобы посмотреть, что в ней упущено. А упущено в ней главное — сознание.

Сёрль приводит 7 положений, которые, по его мнению, составляют основу современных



«ортодоксальных» форм материализма, не оставляющих пространства для сознания. Приведем их.

1. Там, где сознание (mind) является предметом научного исследования, оно с его специфическими особенностями играет весьма малую роль...

2. Наука объективна...

3. Поскольку реальность объективна, наилучший метод изучения сознания состоит в принятии объективной позиции или точки зрения с позиции третьего лица...

4. Единственный ответ на эпистемологический вопрос «каким образом мы в состоянии знать о ментальных феноменах другой системы?» состоит в следующем: мы узнаем о них, наблюдая *поведение* системы...

5. Интеллектуальное поведение и каузальные отношения интеллектуального поведения являются в каком-то смысле сущностью ментального...

6. Каждый факт во Вселенной в принципе познаваем и понимаем в предпринимаемых человеком исследованиях...

7. Все существующее в мире является в конечном счете физическим в том смысле, *в каком традиционно понимается физическое*, то есть как противоположное ментальному<sup>1</sup>.

Каждое из этих семи положений объективистской ортодоксии материализма, утверждает Сёрль, является в лучшем случае ложным. Предлагаемая им альтернатива состоит в следу-

<sup>1</sup> Сёрль Дж. Открывая сознание заново. М., 2002. С. 10—11. Оригинал: Searle J.R. *The Rediscovery of Consciousness*. Cambridge MA, 1992. Далее: Сёрль Дж. Открывая сознание заново. С. ...

307

ющем. Поскольку ментальные феномены существенно связаны с сознанием и поскольку сознание является существенно субъективным, *онтология ментального существенным образом является онтологией от первого лица*. И именно в этом направлении, а не в сторону редукции к функциям, надлежит строить исследования. Полувековая атака на сознание с физикалистских позиций с применением компьютерных аналогий или без них, считает он, не смогла ликвидировать асимметрию между отчетом личности о своих внутренних состояниях «от первого лица» и способами, какими описывают эти состояния «с позиции третьего лица». Поэтому задача реабилитации «субъективной онтологии ментального» становится как никогда актуальной.

У многих философов, в том числе у Деннета, говорит он, сложилось нетерпимое отношение к онтологии субъективности. Они считают ее и метафизически и эпистемологически несостоятельной (постулируются «приватные» сущности, утверждается асимметрия между «внутренним» осознанием личностью ее ментальных состояний и «внешним» их описанием другими людьми). Тем не менее это данность и с ней нужно смириться. «Сознание есть биологическое свойство мозга человека и определенных животных. Оно причинно обусловлено нейробиологическими процессами и в той же степени является частью естественного биологического порядка, как и любые другие свойства, вроде фотосинтеза, пищеварения или деления клетки (mitosis)»<sup>1</sup>. Поэтому главная задача состоит в том, чтобы «вернуть сознание в предмет науки в качестве биологического феномена»<sup>2</sup>.

Сёрль полагает, что можно защищать «онтологию субъективности» и в то же время оставаться натуралистом. Свою позицию он называет «биологическим натурализмом». Это значит, что и сознание и интенциональность следует рассматривать как биологические процессы, причинно обусловленные

<sup>1</sup> Сёрль Дж. Открывая сознание заново. С. 99.

<sup>2</sup> Там же. С. 95.

308

процессами нервной системы нижнего уровня мозга, не сводимыми к чему-либо еще<sup>1</sup>. Сознание не есть знание, самосознание, когнитивность, информация. Это не интроспективное лицемерие в самого себя, а состояние бодрствования, живое, качественное ощущение своего присутствия в мире<sup>2</sup>.

Он категорически не согласен с Деннетом, что квалиа не существуют. Осознанность себя всегда качественна, приватна, субъективна. Это и есть само сознание. Это «*сущностные*» признаки

сознания, которые нельзя заменить никакими иными — поведенческими, функциональными, компьютерными или нейрофизиологическими аналогами. Сознание ассоциируется с физическими, не-интенциональными состояниями мозга в силу того, что и то и другое существует как *одно целое*. И это целое объяснимо на основе *принципа каузальности*: ментальные феномены являются свойствами мозга и нервной системы, а происходящие в мозгу процессы — причинами ментальных феноменов.

<sup>1</sup> Там же. С. 21.

<sup>2</sup> Д.И. Дубровский опубликовал резко критическую статью с подробным разбором книги Дж. Сёрля «Открывая сознание заново». («Вопросы философии», 2003. № 7. С. 92—112). С рядом положений его статьи нельзя не согласиться. Убедительны его аргументы, направленные против толкования Сёрлем способов преодоления дуализма и материализма, смешения отношений онтологии и эпистемологии, его понимания возможностей и границ функционализма и др. Возражение вызывает критика серлевского понимания «субъективной реальности». Нам кажется, что под «субъективной реальностью» Сёрль имеет в виду нечто другое, нежели Дубровский (который тоже защищает идею «субъективной реальности», но в своем особом понимании). Имеется в виду не удостоверяемые интроспекцией субъективно личностные переживания, не самосознание и вообще не какое-то знание, которое в принципе можно компьютеризировать, а что-то, что на русском языке лучше всего назвать «осознанностью», ощущением своего живого бытия или живого присутствия в мире и что дается только «с позиции первого лица» (в отличие, скажем, от компьютера, которому это не дано).

309

Общий вывод Сёрля состоит в следующем. Несмотря на интересные факты, добытые в науках об искусственном интеллекте и в нейрофизиологии о сознательной деятельности, прорыва, требующего радикального пересмотра представления о сознании, не произошло. Мы еще очень мало знаем о функционировании человеческого мозга, его фабрика еще во многом окутана тайной. Можно называть эту ситуацию «кризисом». Выход из него, как и Нагель, он видит в пересмотре стандартного представления об объективизме науки: поскольку спецификой сознания является ее субъективная онтология, нормальная, то есть рационалистически ориентированная, наука должна включить в себя субъективную онтологию, если и не «вовнутрь», то «наряду» с объективной по отношению ко всем другим вещам онтологией.

В рецензии на книгу Сёрля «Открывая сознание заново» Деннет пишет, что автор представил хорошую экспозицию принципов материализма, и ее можно было бы принять за исключением первого пункта, т.к. он не соответствует действительности: на самом деле сознание играет огромную роль в функционализме. Разница состоит в том, что функционализм считает его выводимым ментальным феноменом, а не изначальным базисом интенциональности и сознательности. Главный же минус состоит в том, что «субъективную онтологию» невозможно состыковать с наукой. Задача, стоящая и перед философом, и перед когнитивным ученым состоит не в том, чтобы элиминировать субъективность и забыть о ней, а в том, чтобы найти не противоречащий науке способ ее описания. Именно таков смысл гетерофеноменологии. Она рисует *«нейтральный»* путь, ведущий от объективной физической науки с ее «третьеличностным» методом к методу феноменологического описания, который может (в принципе) отдавать должное наиболее приватным и невыразимым субъективным переживаниям опыта, не забывая о методологических принципах науки<sup>1</sup>. Что касается декларируемого Сёрлем кризиса, то он

<sup>1</sup> Dennett D. *Consciousness Explained*. L., N.Y. P. 72.

310

не является тупиком. «Ортодоксия (по меньшей мере, моя версия) говорит, что кризис совсем не плохая вещь, поскольку, будучи подвергнутой анализу, столь превозносимая "существенная" и "нередуцируемая" субъективность сознания получает разрешение. Вполне возможно иметь и наше сознание, и объективную науку — ценой отказа от нескольких изначальными дезориентирующих контринтуитивий (повседневной психологии и «Картезианского Ока» — *Н. Ю.*). Сёрль же утверждает, что все обстоит наоборот, что мы можем иметь и сознание, и здравый смысл<sup>1</sup>.

## 2. Когнитивизм компьютера и человеческое понимание. Синтаксис и семантика

В 1980 году Сёрль опубликовал статью «Мозг, сознание и программы», в которой представил ряд аргументов, нацеленных охладить чрезмерный энтузиазм философов найти легкое объяснение феномену сознания с помощью компьютерных аналогий. Сёрль имел в виду специальные работы Р. Шенка и Р.Абельсона (1977), Т. Винограда (1973), Дж. Вейценбаума (1965), которые во многом задали подхваченную философами парадигму компьютерного подхода к сознанию. Хотя имя Деннета в ней не упомянуто, он, несомненно, целился и в его адрес. В последующих книгах, особенно в книге «Тайна сознания» (1998), он выходит на прямую полемику с Деннетом.

Сёрль проводит различие между позициями «сильного AI» и «слабого AI» (или «осторожного»). Согласно «слабому AI», основная ценность компьютера в исследовании сознания инструментальна; сравнение естественного и искусственного интеллекта дает возможность более строгом и точным образом формулировать и проверять гипотезы. Согласно «сильно-

<sup>1</sup> Journal of Philosophy. Op. cit. P. 196.

311

му AI», компьютер — не просто инструмент в исследовании сознания; соответственно запрограммированный, он представляет собой некое *сознание* в том смысле, что способен *понимать* и иметь другие присущие человеку когнитивные состояния. У Сёрля нет возражений против притязаний «слабого AI», он сам является его сторонником, притязания же «сильного AI» он считает безосновательными и дезориентирующими.

Никто не спорит, говорит он, что использование компьютерных программ для моделирования процессов сознания явилось большим прорывом в исследовании работы человеческого интеллекта. Модели — мощные инструменты в познании любого объекта. Но только в том случае, если не принимать моделирование за дублирование. Мы моделируем, например, ураганы, но модели ураганов не становятся от этого ураганами. Это относится и к компьютерам. Как бы успешно компьютер ни моделировал когнитивные процессы, он остается только машиной. Главное его отличие от человека состоит в том, что он *не понимает смысла того, что он делает*. Не понимает обычные рассказы, в которых, как правило, содержатся подтексты, двусмысленности, коннотации и др. Многие сторонники сильного AI выдвигают такое возражение: все дело в несовершенстве программ, если же создать программы, насыщенные максимальной информацией, машина смоделирует человеческие способности, в том числе *способность понимать рассказы и релевантно отвечать на относящиеся к ним вопросы*. Сёрль отвечает на него так: компьютерная программа представляет собой только связку символов, организованную по определенным синтаксическим правилам. Сами по себе эти символы лишены семантических значений; для их семантического наполнения требуется «означиватель» — субъект, чья означивающая способность коренится в его ментальности.

Для доказательства защищаемого им тезиса о неправомерности отождествления когнитивных операций человека с операциями компьютера, прежде всего тезиса о неспособности AI к *пониманию*, он представил мысленный эксперимент, извест-

312

ный как «*Аргумент Китайской Комнаты*». Аргумент получил широкий отклик, в том числе и критический. В ответ Сёрль несколько раз улучшал его, но мы не будем останавливаться на его модификациях и представим его в первоначальном виде, с которым и полемизировал Деннет.

Аргумент заключается в следующем. Рассматривается воображаемая ситуация, когда в комнате находится англоговорящий человек-оператор (или машина), не владеющий китайским языком. Человека познакомили с рядом написанных по-английски правил трансформации китайского языка, которые он применяет для обработки поступающих к нему сообщений на китайском языке, манипулируя с ними в соответствии с этими правилами. Результатом его манипуляций с иероглифами является производство новых сообщений по-китайски. Получая их, находящиеся вне комнаты люди, говорящие на китайском языке, в состоянии понять весь этот обмен сообщениями

как вполне умопостигаемый разговор (с точки зрения теста Тьюринга он вполне успешен). Сидящий же в комнате человек, манипулирующий китайскими иероглифами в соответствии с правилами китайского языка (то есть программой) чисто формальным путем, не будет иметь ни малейшего представления относительно содержания разговора<sup>1</sup>.

Суть аналогии с Китайской Комнатой состоит в предположении, что воображаемый человек-оператор, находящийся в комнате (или машина вместо него), хотя и посылает разумные сообщения (outputs), просто осуществляет синтаксические трансформации с входящими сообщениями (inputs), которые для него самого только бессмысленные закорючки. Компьютер, говорит Сёрль, ведет себя точно так же, как и человек-оператор в Китайской Комнате; он совершает вычислительные операции с помощью формально определенных элементов. Отсюда следует вывод: поскольку эффект компьютера не более чем синтаксическая манипуляция с символами-закорючками, у него не мо-

<sup>1</sup> Сёрль Д. Мозг, сознание и программы // Аналитическая философия: становление и развитие. М., 1998. С. 378.

313

жет быть интеллекта или понимания. «Все, что имеется у компьютера,— это символы и еще раз символы»<sup>1</sup>.

Важно подчеркнуть, что смысл предложенного Сёрлем мысленного эксперимента с Китайской Комнатой состоит не просто в доказательстве, что синтаксический уровень сам по себе никогда не превратится в семантический и что запрограммированный цифровой компьютер с правильными входом и выходом не может быть сознанием точно в таком же смысле, в каком сознание присуще человеку. Главная его задача — показать неустранимость субъекта или «означивателя» из реального процесса понимания. Отсюда следует, что всякое объяснение сознания не может ограничиваться «внешним» — компьютерным или нейрофизиологическим — уровнем описания и должно включать в себя «внутреннее», то есть субъективность. «Дистинкция «ментальное» — «нементальное» не может зависеть только от точки зрения внешнего наблюдателя. Она должна быть внутренне присущей самим системам; иначе любой наблюдатель имел бы право, если бы пожелал, трактовать людей как нементальные феномены, а, скажем, ураганы — как ментальные»<sup>2</sup>. Сознание в полном смысле этого слова связано только с органическим носителем — мозгом — и объяснение этому феномену следует искать только в его особенностях.

В книге «Открывая сознание заново» (1992) Сёрль повторил «Аргумент Китайской Комнаты», добавив к нему новые положения. Не только семантика, говорит он, но и синтаксис задаются наблюдателем. Заявление сторонников «сильного AI», что мозг является цифровым компьютером, совершенно ложно. Ничто на свете не является внутренне вычислительным устройством — ни физические вещи, ни мозг; вычисления осуществляют только наделенные интенциональностью сознательные агенты, и, если вам нравится, их можно называть «гомункулами». «Исходя из определения вычислений и позна-

<sup>1</sup> Там же. С. 397.

<sup>2</sup> Там же. С. 385.

314

ния, вычислительная наука никоим образом не сможет быть естественной наукой, потому как вычисление не является внутренне присущим свойством мира. Оно приписывается в зависимости от наблюдателей... <...> ...Без гомункула... у нас даже нет синтаксиса, с которым мы могли бы работать»<sup>1</sup>. Более того, сторонники «сильного AI» и когнитивисты не в состоянии представить *каузальные основания для познания*, ибо без гомункула у компьютера, так же как и у мозга, есть только схемы, а схемы не имеют никаких каузальных воздействий помимо тех, которые присущи выполняющим программы приборам.

Одним из серьезных заблуждений когнитивизма Сёрль считает представление о мозге как переработчике информации. Многие когнитивные ученые думают, что мозг и сознание человека обрабатывают информацию точно таким же образом, как и компьютер, а если — в идеале — у компьютера будет та же самая программа, что и у мозга, мы получим тождественную обработку

информации. Отсюда ими делается вывод, что обработка информации и есть сущность ментального. Это совершенно неверное представление о когнитивной работе мозга. Когда компьютеру дается задание вычислить, сколько будет 2+2, он напечатает «4». В отличие от человека у него нет никакого представления о том, что знак «4» означает число 4 и что вообще означает четыре. И дело не в отсутствии у него какой-то дополнительной информации, а в том, что его символы лишены интерпретации и *интенциональности*, входящих в человеческую обработку информации. «Что касается внутренне присущих мозгу операций, то он (мозг — *Н. Ю.*) не осуществляет никакой переработки информации. Он является особым биологическим органом, и особые нейрофизиологические процессы в нем вызывают особые формы интенциональности»<sup>2</sup>. Несмотря на влияние социокультурных факторов, сознание, настаивает Сёрль, представляет собой

<sup>1</sup> Сёрль Дж. Открывая сознание заново. С. 196, 198.

<sup>2</sup> Там же. С. 208.

315

прежде всего продукт биологической эволюции и должно быть понято и объяснено исходя из природных каузальных связей и прежде всего из функционирования биологической материи — мозга. Существуют грубые, слепые нейрофизиологические процессы и существует сознание, ничего другого кроме этого нет. Ментальные события являются свойствами мозга в таком же смысле, в каком свойство быть жидкостью является свойством воды. Иначе говоря, в нем нет ничего такого, что придумали философы в последние десятилетия: никаких информационных процессов, никакого следования правилам, никакого языка мысли и никакой универсальной грамматики.

Аргумент Китайской Комнаты Сёрля содержит серьезную угрозу стратегии Деннета. Если окажется, что он корректен, тогда следует восстановить в правах и феноменализм, и самость как «Центрального означивателя», и интроспекцию, и все остальное, что он с таким старанием деконструировал. Этого, конечно, он не может допустить и поэтому в своих работах неоднократно обращался к этому аргументу, формулируя все новые и новые возражения.

Одним из контраргументов сторонников сильного AI, выдвинутого против тезиса Сёрля о неспособности компьютера к пониманию, является так называемый *системный аргумент*. Суть его состоит в том, что отдельный человек (или машина), запертый в комнате, действительно, не понимает рассказа, но он представляет собой лишь часть некоей системы, а система в целом понимает рассказ. Деннет является сторонником именно системного подхода. Для того чтобы показать его работоспособность, он придумал свой мысленный эксперимент — «Диалог судьи с Китайской Комнатой». Эксперимент должен доказать, каким образом Китайская Комната в целом (или беспредельно гигантская компьютерная программа) сможет успешно пройти «тест Тьюринга».

«Судья: Слышали ли вы об ирландце, нашедшем чудесную лампу? Когда он потерял ее, появился Джинн и пообещал выполнить три желания. «Я хочу pintу пива», молвил ирландец, и тотчас же появилась кружка с пивом. Ирландец с

316

жадностью стал сначала смаковать его, а затем пить большими глотками, однако уровень пива в кружке всегда чудесно восстанавливался. Через какое-то время Джинн стал проявлять нетерпение. «Ну ладно, а как насчет твоего второго желания?» На что ирландец в паузе между двумя большими глотками ответил: «Я думаю, что было бы недурно для меня получить еще одну кружку пива».

Китайская Комната: Очень смешно. Нет, я впервые услышала анекдот, но вы же знаете, что я считаю этнические анекдоты проявлением дурного вкуса. Хотя я и посмеялась, на самом деле мне кажется, что вам следовало бы найти другую тему для нашей дискуссии.

Судья: Замечание вполне справедливое, однако я рассказал анекдот для того, чтобы вы объяснили его мне.

Китайская Комната: Как это скучно! Никогда не следует объяснять шутки.

Судья: Тем не менее это мой вопрос в тестировании. Можете ли вы объяснить, как и почему



«работают» шутки?

Китайская Комната: Ну, если вы настаиваете. Понимаете ли, все здесь зависит от допущения, что чудесно заполняющаяся кружка будет заполняться вечно, поэтому в распоряжении ирландца есть все пиво, которое он когда-либо в состоянии будет выпить. Посему у него едва ли был резон хотеть еще одну кружку, но он настолько глуп (это как раз то, против чего я протестую) или потерял рассудок от алкоголя, что не понимает это допущение и поэтому, не раздумывая, желая продлить свое удовольствие от правильности его первого желания, заказал вторую кружку. Эти подспудные допущения не являются, конечно, истинными, но они и составляют часть особого двусмысленного знания в рассказывании шуток, в котором мы на время воздерживаемся от нашего неверия в чудеса и тому подобное. Между прочим, мы могли бы вообразить какое-нибудь продолжение истории, в которой ирландец в конце концов оказался бы «прав» в его втором желании. А может быть, он планировал устроить большой прием гостей, и одна кружка пива не смогла бы заполняться

317

достаточно быстро, чтобы удовлетворить всех жаждущих (заготавливать пиво заранее не имеет смысла, все мы знаем, что постояя оно теряет свой вкус). Нам свойственна манера не думать о таких сложностях, чем частично объясняется, почему анекдоты срабатывают. Этого достаточно?»<sup>1</sup> Программа, которая могла бы на машинно-виртуальном уровне генерировать разговор Китайской Комнаты и разумные ответы на вопросы судьи, выглядела бы, по мысли Деннета, следующим образом. Приступив к грамматическому разбору первых слов «Слышали ли вы о...», некоторые из квази-понимающих homunculi, отслеживающих анекдоты, были бы приведены в состояние активности, в результате которой из компьютерной памяти будет извлечено множество стратегий, имеющих дело с воображаемыми и фиктивными объектами. Когда пришел черед грамматического разбора слов «чудесная лампа», программа была бы подготовлена к тому, чтобы низко оценить приоритетность истинностных ответов, объясняющих, что такой вещи, как чудесная лампа, не существует. Будут приведены в действие квази-понимающие homunculi (или «демоны»), сенситивные к негативным коннотациям в рассказывании этнических анекдотов, постепенно подводя ко второй теме первого ответа Китайской Комнаты... И так далее. О том, что такой процесс далеко не фантазия, говорит Деннет, свидетельствует достаточно успешная практическая инсталляция в программах стандартных нарративных образцов анекдотов с джиннами (М. Минский, 1975) или рукописных текстов (Р. Шэнк и Р. Эбельсон, 1977).

Конечно, поясняет Деннет, программа, способная к подлинному пониманию анекдота, должна быть в высшей степени многоуровневой системой, максимально насыщенной информацией о мире. Предполагается, что она должна обладать и знанием, и мета-знанием, и мета-мета-знанием о своих собственных ответах, и знанием о предполагаемых ответах собеседника, и знанием ее собственных мотиваций, и мотиваций

<sup>1</sup> Dennett D. *Consciousness Explained*. L., N.Y. P. 437.

318

собеседника, и еще многим и многим другим. «Может быть, продуцирование миллионов действий всех этих высоко структурированных частей, в конце концов, и придаст системе свойство подлинного понимания»<sup>1</sup>.

Ложность стратегии Сёрля Деннет видит в смазывании существенного различия между возможностями простых и гигантских компьютерных систем. Мысленный эксперимент с Китайской Комнатой принимает за образец слишком элементарный случай. В простых системах, действительно, нет ничего даже отдаленно похожего на подлинное понимание, имеет место просто бессознательная рутина трансформации одних веревок символов в другие веревки в соответствии с некоторыми механическими или синтаксическими правилами. В гигантских системах Искусственного Интеллекта дело с пониманием может обстоять принципиально по-иному. Они состоят из множества распределенных в них «глупых квази-понимающих» элементов. Взаимодействуя в огромной и чрезвычайно сложной системе, в целом они могут продуцировать понимание. Такая «Комната» выдержала бы тест Тьюринга.

Сёрль, конечно, возразит, что большее количество непонимающего, неважно насколько оно больше, никогда не добавит подлинного понимания. Многим его возражение кажется бесспорным. Тем не менее, говорит Деннет, оно ошибочно. За ним скрывается трудноустраняемая картезианская иллюзия о невозможности подлинного понимания без «добавки» — души, то есть иллюзия дуалиста. «Если мы, материалисты, убежденные, что для достижения понимания наши мозги так или иначе обходятся своими собственными ресурсами, не нуждаясь в помощи чуда, мы должны также допустить, что подлинное понимание каким-то образом достигается в процессе интеракций между огромным количеством подсистем, ни одна из которых сама по себе не понимает вещь. Аргумент, начинающийся с того, что «этот маленький кусочек активности мозга не понимает китайский язык, и его не в состоянии понять

<sup>1</sup> Ibid. P. 438.

319

большая часть, куда входит этот кусочек...», венчается нежелательным выводом, что даже активность всего мозга недостаточна для объяснения понимания китайского языка<sup>1</sup>. Конечно, замечает Деннет, вообразить, каким образом большее количество «непонимающего» может в итоге дать понимание, трудно. Однако еще труднее вообразить, каким образом ничего не понимающий кусок органической материи — мозг — продуцирует язык и мысли.

Если посмотреть на сознание как на состоящее из постоянно меняющихся и перекрывающихся потоков когнитивных процессов (т.е. Модели Множественных Набросков), можно предположить, что между присутствием и отсутствием сознания граница размыта. Когда одна функция громоздится на другую, а та на последующую, создавая чрезвычайно сложную систему функциональных взаимоотношений, наша интуиция о такой системе, а вместе с ней и «субъективная онтология» Сёрля, оказывается беспомощной. Конечно, *ни один* отдельный нейрон не является сознательным; не становится он сознательным и при вступлении в связь с другим отдельным нейроном. Когда же связи нейронов повторяются миллиарды раз, в конце концов получается сознание. Одним словом, если исходить из того, что сознание есть свойство мозга как функционально целого, неразумно искать его в частях, составляющих это целое, — независимо от того, рассматриваем ли мы мозг как сложную систему нейронов или как сложную систему функций. Тот же самый аргумент применим и к роботам-компьютерам: одна вычислительная функция не является сознательной, бесконечное же добавление одних кусочков вычислительных функций к другим в итоге может дать осознающего робота<sup>2</sup>.

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 438—439.

<sup>2</sup> Некоторые когнитивные ученые считают используемый Деннетом аргумент в отношении компьютеров слабее, нежели в отношении людей. Кэвин Корб, например, говорит, что в случае мозга мы наперед знаем, что ему присуще сознание, и это создает уверенность, что система нейронных функциональных сложностей в принципе производит что-то осознанное. Относительно же сознания робота мы ничего не знаем наперед. Например, наши прыжки могут становиться все выше и выше по сравнению с предшествующими, однако из этого не следует, что после какого-то неопределенно большого количества прыжков мы достигнем Луны. (Korb V.K. *Stage Effects in the Cartesian Theatre: A Review of Daniel Dennett's Consciousness Explained* // *Psyche*. № 1 (4). Dec. 1993.)

320

Как бы то ни было, Деннет твердо убежден в большей продуктивности его системного аргумента и всей его стратегии, нежели в скептической стратегии Сёрля. Он видит в ней потенциал для преодоления заключенной в дуализме опасности бесконечного регресса, не прибегая при этом к «сущностям», к «конечным основаниям» или к «трансцендентным тайнам». В ответ на упреки, что в его изображении сознание выглядит порцией текста без агента, которому другой текст придает значимость и смыслы, Деннет говорит, что удаление субъекта (агента) — единственный путь выпрыгнуть из порочного круга, создаваемого идеями субъективной онтологии и «агента как означивателя».

\* \* \*

В заключение нашей экспликации полемики между Деннетом и Сёрлем хотелось бы сказать

следующее. Конечно, Сёрль не в меньшей мере, чем Деннет, является материалистом, объективистом, рационалистом и эволюционистом. Его самоидентификация с помощью термина «биологический натурализм» не меняет дела. Как и Деннет, он стремится разгадать загадку сознания, ориентируясь на науку; в отличие от Нагеля, считающего игру с компьютерными метафорами пустой тратой времени, сравнение естественного и искусственного интеллекта у него — важный инструмент исследовательской стратегии. Не меньше, чем Деннет, он хочет нарисовать монистическую картину мира, обойдя ловушки картезианского дуализма. Самое интересное состоит в том, что каждый из

321

них видит в стратегии оппонента сползание в дуализм. Деннет упрекает Сёрля за то, что его онтология субъективного ведет к реставрации картезианского дуализма, а тот, в свою очередь, упрекает Деннета в изобретении новой версии дуализма — компьютерных программ и органического мозга.

Сопоставляя исследовательские стратегии этих двух философов, нам представляется, что они символизируют оппозиции, сложившиеся в современной философии сознания между биологическим, социобиологическим и социолингвистическим подходами к объяснению сознания. У Сёрля — биологический крен: разгадку тайны сознания он ищет в особом биологическом устройстве мозга. В то же время, отрицая причинно-следственные отношения между мозгом и сознанием и рассматривая сознание как новое эмерджентное качество в эволюционном развитии, он связывает его объяснение с переопределением каузальности. В отличие от Сёрля, Деннет помещает сознание на стыке биологического и социального и связывает его с влиянием социального на биологию мозга; если же начинать с биологии и нейрофизиологии мозга, мы никогда не найдем пути к объяснению сознания. Отсюда разная фокусировка свойств сознания. Если Деннет делает акцент на когнитивной составляющей сознания и паттернах языка и культуры, Сёрль в первую очередь сосредоточен на биологических привязках сознания и лишь потом на социальных.

В нашей оценке этих двух стратегий есть одно важное соображение. Биология мозга еще слабо изучена, мозг состоит из миллионов перекрещивающихся связей и функций, компьютерная симуляция которых маловероятна. Искать на этом пути тайну сознания — значит, отодвигать ее разгадку в бесконечно далекое будущее. Другое дело когнитивная составляющая сознания; в ней тоже много белых пятен, но она лучше поддается исследованию и компьютеризации. Последнее обстоятельство во многом определило более интенсивное развитие именно информационно-процессуального (функционального) уровня анализа.

Последнее замечание относится к философской и научной составляющей спора. И Деннет и Сёрль хорошо осведомлены о том, как обстоят дела в когнитивных науках и исследованиях по AI. Они не страдают самомнением и не полагают, что их собственную работу следует расценивать как научную или почти научную; они считают ее философской, т.е. содержащей существенный элемент спекулятивности. Деннет предупреждает читателя, что его видение решения проблем во многом основано на его собственных интуитивных импульсах (intuition pumps). Поэтому его рассуждения не следует рассматривать как прямолинейные демонстрации фактов, «они больше являются искусством, нежели наукой»<sup>1</sup>. Если в них содержатся перспективы, которые могут быть исследованы систематическими научными методами, я посчитаю это большим достижением. Если же они направляют по ложному пути — это достойно сожаления. Ибо даже хорошие инструменты могут использоваться неправильно.

<sup>1</sup> Dennett D.C. *Consciousness Explained*. P. 440.

322

#### IV. Деннет versus Рорти: МЕТАФОРИЗМ, РЕАЛИЗМ И «БАЗИСНЫЙ ИНСТИНКТ»

Будучи неисправимым узко мыслящим и чуждым историчности аналитическим философом, я всегда ищу удобное оправдание тому факту, что не читал Гегеля, или Хайдеггера, или Деррида, или каких-либо других парней, которые не снизили до любезности думать по-английски.

*Дэниел Деннет*

Там, где Деннет видит шкалу, на одном конце которой истеричный реализм, на другом — мой собственный истеричный, хотя и модный антиреализм, а себя в благоразумной — не истеричной — середине, я вижу достаточно жесткий раскол между людьми, которые хотят видеть в естественной науке исключительность, отличную от всего остального в высокой культуре, и людьми вроде меня, которые смотрят на нее как на часть разговора, служащего только определенным целям.

*Ричард Рорти*

Томас Нагель и Джон Сёрль — это оппоненты Деннета, так сказать, «справа» — со стороны защитников ментальности и субъективности. Оппонентами «слева» являются элиминативные материалисты, предлагающие сдать ментальный словарь в архив. Одним из них является Ричард Рорти. В данном разделе мы рассмотрим полемику между Деннетом и Рорти.

324

Для целей нашей работы рассказ о ней полезен прежде всего потому, что в ней более ясно представлены эпистемологические принципы философии Деннета (понимание реализма, инструментализма, истины и др.) и его общие рационалистические воззрения. Рассказ поучителен и для понимания разницы в философско-идеологических установках этих двух мыслителей. Как уже говорилось, Деннет работает в рамках социобиологической парадигмы. Ее специфика лучше видна в сопоставлении с «историцизмом» Рорти и социолингвистической парадигмой, характерной для постмодернистского варианта прагматизма. Думается, что рассказ о полемике небезынтересен для российского читателя и по другой причине: ее содержание позволяет судить о связи между «твердым ядром» философии, куда входят, в первую очередь, проблемы сознания и реализма, и общей философской идеологией. К тому же, как нам кажется, показ этой связи поможет скорректировать одностороннее впечатление, сложившееся в российской литературе о расстановке сил в конфликте между сторонниками теоретизма философского знания и постмодернистами, сводящими философию к одному из жанров «литературы».

Сначала немного истории. Начало творческих исканий Деннета и Рорти падает на период 60—70-х годов, во многом революционный для англоязычной философии, когда под влиянием неопозитивизма и лингвистического анализа все области философского знания были захвачены поисками новых путей работы. Дети перестройки того времени Рорти и Деннет (так же как П. Фейерабенд, Х. Патнэм, Т. Нагель, Д. Сёрль, П. Черчленд, Д. Дэвидсон, Дж. Фодор и др.) оказались в атмосфере жарких споров о новом понимании науки и философии, рациональности, эмпирического и теоретического, языка, сознания. Больше всего их привлекли баталии, развернувшиеся вокруг последней проблемы. К тому времени за плечами Рорти была нашумевшая статья «Проблема духовного и телесного, приватность и категории» (1965), обширное предисловие к антологии «Лингвистический поворот» (1967), статья «Не-

325

исправимость как признак ментального» (1970). Деннет только что опубликовал книгу «Содержание и сознание» (1969). Она тут же попала в поле зрения Рорти; в молодом авторе (Деннет на 11 лет моложе Рорти) он увидел единомышленника и продолжателя терапевтической стратегии Витгенштейна— Райла на очищение философии сознания от ее «корневой метафоры» — декартовского «Внутреннего Ока».

Пolemика долгое время не выходила за рамки внутрицеховых для аналитической философии споров и носила, говоря словами Деннета, характер «конструктивного несогласия» (касалась уточнения некоторых аспектов проблемы сознания). Спустя несколько десятилетий, в 2002 году, в сборнике, подготовленном к 70-летию Рорти («Рорти и его критики», 2000), Деннет

признает, что ретроспективно ему стало виднее, что еще в начале его работы он неосознанно усвоил ряд «посланий» и понятий, содержащихся в ранних статьях Рорти, — «привилегированный доступ к собственному сознанию», «неисправимость (incorrigibility) ментального» и других, которые укрепили его в мысли о необходимости полномасштабной критики ментализма. Смысл введенного Рорти термина «неисправимость ментального» означал не просто, что ментальные отчеты не поддаются проверке на ошибочность, а то, что они в принципе не могут быть опровергнуты, поэтому их референциальность должна быть подвергнута отрицанию. Вопреки его более поздним «прагматистским» взглядам на философию, Рорти, считает Деннет, не осознавая, «внес вклад в науку — в наше объективное знание того способа, каким существует и мир и сознание, в философское устремление, от которого он затем отрекся. Его собственный «разговор» оказался больше, чем просто разговор»<sup>1</sup>.

В 70-е годы развитие их взглядов пошло по разным руслам. По мере очарованности Рорти идеями Ф. Ницше, М. Хай-

<sup>1</sup> Dennett D. The Case for Rorts // Rorty and his Critics. Ed. by R.B. Brandom. Oxford, Melden MA, 2002. P. 92.

326

деггера, Г.-Г. Гадамера, Ж.Деррида,<sup>1</sup> а Деннета — все большей вовлеченности в проблематику когнитивных наук, искусственного интеллекта, биологии и роста интереса к идеям социобиологов (Э. Уилсона, Р. Докинса и др.), пространство несогласия становилось все шире. Взятый Рорти ориентир на историцизм и «философию-как-литературу» все меньше пересекался с ориентиром Деннета на «философию-как-науку». Хотя на самом деле Рорти никогда не обрывал окончательно корни, связывающие его с взрастившей его аналитической философией, продолжая следить за тем, что делается в ней, и выискивая там теоретические аргументы в пользу постмодернистского варианта прагматизма. Несмотря на поношения аналитиков, он любит — в качестве *infant terrible* — участвовать в их дискуссиях. Как бы то ни было, к концу XX века два мыслителя оказались по разные стороны баррикады в понимании философской идеологии: Рорти — в стане критиков аналитической философии и защитников постмодернистского варианта прагматизма, Деннет — в стане защитников теоретизма философского знания, рационализма, объективизма, истины.

И Деннет и Рорти — не только оригинальные мыслители. Они известны как философские стилисты, изобретательно работающие с метафорами и аналогиями, обыгрывающие примеры и виртуальные ситуации. Вместе с тем по выбранному жанру и характеру мышления они существенно разные. Рорти — историцист, рисующий картины крупными мазками на широком фоне современной философской ментальности, у

<sup>1</sup> В отличие от Деннета, получившего в студенческие годы «куайновскую» натуралистическую закалку, Рорти прошел выучку в Йельском университете, где слушал лекции Ричарда МакКеона, который привил ему вкус к метафилософии, к тому, что называется «взглядом с птичьего полета». В основном творчество Рорти представляет собой попытку соединить собственно аналитическую работу с метаментальностью, то есть с рефлексивной ее оценкой в рамках современного историко-философского контекста.

327

него всегда под рукой исторические параллели и аналогии. Деннет — аналитик, предпочитающий «инженерный» подход к проблемам и настороженно относящийся к обобщающим констатациям (всегда находится «черный лебедь», опровергающий общее суждение).

Пикантность ситуации состоит в том, что Деннет и Рорти — друзья и во многом единомышленники. Рорти помещает себя в одну с Деннетом традицию философии сознания (которую он именует «традицией Райла—Деннета»). Поделив философов на «героев» и «антигероев», он без колебания поместил Деннета в почетную когорту «героев» (наряду с У. Куайном, Н. Гудменом, У. Селларсом и другими аналитиками), на работах которых, по его собственному признанию, он «паразитирует». У Деннета не менее высокая оценка творчества коллеги: его метафилософский талант, умение работать с контекстом и сопоставлять позиции, историко-философская эрудиция будят воображение и мысли; Рорти вполне заслужил ту широкую читательскую аудиторию, которую он имеет в мире.



Из представленного в предыдущих главах содержания книги читатель может судить, что сам Деннет далеко не безгрешен по части деконструктивизма — разрушения несущих конструкций традиционного здания философии, и похвалы Рорти вовсе не безосновательны. Сознание объясняется таким образом, что от него ничего не остается (во всяком случае, в прежнем его понимании). Тенденция к радикализму, как мы показали, имеется и в эпистемических воззрениях Деннета; реляционизм и интерпретивизм трудно соединяются с традиционным реализмом. Хотя здесь он более осторожен и, не желая оказаться в стане инструменталистов и ирреалистов, выстраивает сложные (эволюционистские) конструкции. Будучи опытным метафилософом, Рорти подмечает половинчатость Деннета и логические цепочки, связывающие его антиэссенциалистские интенции (и аналитической философии) с аналогичными интенциями Ницше, Хайдеггера, Деррида и др. Во всяком случае, Рорти не без основания предлагает Деннету протянуть руку дружбы постмодернизму. Деннет все же держит ее в кармане.

328

## Рорти о Деннете

### 1. «Коперниканский переворот» в философии сознания: от эссенциализма к прагматизму

Несколько слов о деконструктивистском проекте самого Рорти. У него три основные задачи. Во-первых, низложить трансцендентность во всех ее видах («сущности» Аристотеля, «Бог» религиозной философии, «абсолютная идея» Гегеля, «вещь в себе» Канта, «дух» картезианцев, «онтология субъективности» Сёрля и др.). Во-вторых, развенчать авторитет платоновско-декартовско-кантовской гносеологической традиции с ее претензиями на реализм, «зеркальность», «проникновение за занавес явлений», «приближение к истине». В-третьих, положить конец мифу о «естественности» философских проблем и «теоретичности» философской работы. Основными посылками в реализации его деконструктивистского проекта являются две максимы Витгенштейна: «границы языка являются границами мысли», все философские теории представляют собой «языковые игры». То есть весь разговор на философские темы Рорти переносит в социолингвистическую и прагматическую плоскость.

Поскольку наиболее упорно сопротивляющимся «бастионом трансцендентности» являются концепции сознания, настаивающие на существовании сознания как особого внутреннего феномена субъекта, требуется их радикальное опровержение. Деннет привлекает Рорти прежде всего своей решительностью покончить с этим бастионом при помощи «экстерналистского» гетерофеноменологического метода. (Не последнюю роль играет и авторитет Деннета в области когнитивных наук.) Тактический замысел Рорти состоит в том, чтобы использовать идеи Деннета в виде критического тарана против менталистских позиций и одновременно указать на непоследовательность его атак на трансцендентность и уговорить занять более жесткую позицию по отношению к противникам (Нагелю,

329

Сёрлю и др.). Битва за союзника — Деннета — и призыв его к большей жесткости имеют для Рорти принципиальное значение: в отстаивании первыми «внутреннего пространства» сознания и в непоследовательности второго он видит угрозу максиме Витгенштейна «границы языка являются границами мысли», составляющей основной каркас его прагматизма.

Одна из главных констатаций книги Рорти «Философия и зеркало природы» и других книг<sup>1</sup> состоит в следующем. В философии последнего времени наметилось движение к радикальному пересмотру пространства сознания, аналогичное тому, которое было в науке, когда Галилей и Ньютон подвергли радикальному пересмотру физическое пространство: сознание выводится из внутренней сферы вовне, в социальность, и перестает играть роль центрального светила, в соотношении с которым рассматриваются все остальные проявления человека. Отказ от центрального положения сознания сравним с «коперниканским переворотом», когда люди осознали,

что не Солнце движется вокруг Земли, а Земля вокруг Солнца. Переворот, в котором участвуют операционисты, прагматисты, Витгенштейн, Райл, Гудмен, постмодернисты, состоит прежде всего в вытеснении реликтов аристотелизма из философии сознания — сущностных вопросов о природе сознания — и в принятии *номинализма*. В свете этой революции следует оценивать и философию Деннета<sup>2</sup>.

Основное и главное достоинство творчества Деннета Рорти видит в продолжении терапевтической стратегии Витгенштейна— Райла по освобождению языка философии от не имеющих референтов химер. А именно в том, что он низложил, прибегнув к удачной метафоре «центр нарративной гравитации», миф о наличии у нас метафизической самости и заменил «корневую

<sup>1</sup> Rorty R. *Philosophy and the Mirror of Nature*. Princeton, 1979; Rorty R. *Consequences of Pragmatism*. Minneapolis, 1982; Rorty R. *Contingency, Irony and Solidarity*. Cambridge, N.Y., Sydney, 1989.

<sup>2</sup> Rorty R. *Contemporary Philosophy of Mind // Synthese*. V. 53. № 2. November 1982. P. 337—343.

330

метафору» всей прежней философии — «сознание» — на «верование в сознание». Его «гетерофеноменологический» метод — это прежде всего лингвистический метод, а интенциональная установка в его руках является способом интерпретации лингвистического поведения организмов, которым приписывается сознание. Нагель, Сёрль и другие философы, обвиняющие Деннета в устранении сознания из нашей картины мира, не осознают глубины и последствий переворота, совершенного в традиции Витгенштейна—Райла—Деннета, последствием которого явился полный пересмотр «Вселенной Сознания».

Деннет, считает Рорти, представил убедительные доказательства, почему в рамках фолк-психологии нам *кажется*, что существует внутренний феноменальный мир, почему нам *представляется*, что существует различие между чем-то, что кажется нам розовым, и что реально видится розовым, и почему все эти кажимости не следует принимать за реальное. После выхода в свет книги «Сознание объясненное» (1991) многие критики обвинили его в «зомбизме», в том, что он обходит суть сознания — а именно тот факт, что сознание лежит вне досягаемости языка и культуры и может познаваться только изнутри. Нагель назвал попытки Деннета и других философов освободиться от субъективности симптомом детского восстания «против самого философского импульса»<sup>1</sup>. Все эти критики, по мнению Рорти, не осознавая глубины и последствий происшедшего в философии сознания сдвига, пятаются назад к аристотелевскому взгляду на мир и познание.

Корни расхождения прагматических номиналистов (к которым он относит Деннета и себя) и трансценденталистов (в число которых он помещает Нагеля и Сёрля) Рорти видит прежде всего в разном понимании *целей философской работы* и разной трактовке *смысла философского объяснения*. Для прагматистов объяснить сознание — значит исследовать все его отношения и

<sup>1</sup> Цит по: Rorty R. *Holism, Intrinsicity and the Ambition of Transcendence // Dennett and His Critics. Demystifying Mind*. Ed. by Bo Dahlbom. Cambridge MA., 1993. P. 189.

331

функции с целью практических *предсказаний и установления контроля* над интеллектуальным поведением. Нагель и Сёрль считают предсказание и контроль внешних проявлений сознания недостаточными для его понимания, они требуют объяснения его как определенного феноменального свойства субъекта. Последнее предполагает обращение к внутреннему, до-лингвистическому, трансцендентному пространству, судить о котором можно только на основе субъективного опыта «от первого лица».

Рорти убежден, что, двигаясь с разных сторон, Витгенштейн, Райл, Дэвидсон, Гудмен, Деннет нанесли сокрушительный удар по идее «привилегированного доступа к своему собственному сознанию». После него должно рухнуть и картезианское представление о сознании как «Оке», лицезреющем самое себя. То, что образ «Ока» оказался устойчивым, объясняется не столько доказательной силой защищающих его теоретических аргументов, сколько апелляцией к интуитивным (иллюзорным) представлениям человека о себе. Человеку кажется, что ментальные

термины типа «надеюсь», «мне больно», «хочу» обязательно должны иметь референты в виде особых ментальных состояний сознания.

Различие между «коперниканским» и «аристотелевским» объяснениями объекта Рорти поясняет с помощью аналогии между изучением сознания и изучением движения. В науке наших дней принято считать, что эмпирически исследовать движение тел можно при помощи вычисления изменения их положения во временной последовательности. Тем не менее немало было и есть философов, видящих в таком предположении упрощение, упускающее внутреннюю суть движения. Аристотель, определив движение как «актуализацию потенциала как потенциала», связывал естественное движение с понятием субстанции. Лейбниц полагал, что движение требует не только математически-физического, но и метафизического объяснения. «Аристотелевский» подход к исследованию движения практиковали Бергсон, Джемс, Уайтхед, Вейс и множество других философов, принимающих идею «процесса» в качестве фундаментальной категории. И вот подобно тому как нынешние спекулятивные процесс-философы

332

упрекают ученых за то, что те упускают «чистую сущность движения» — «движение как движение», — Нагель и Сёрль упрекают Деннета за то, что его гетерофеноменология упускает главное: сознание как сознание, его особую реальность. Для прагматистов (к которым Рорти причисляет и Деннета), отказавшихся от идеи абсолютного движения и рассматривающих движение с точки зрения относительности, такой вопрос некорректен. «Философы этого типа не видят смысла в употреблении слова «реальное» по отношению к сознанию»<sup>1</sup>.

Отказавшись от эссенциализма и приняв на вооружение *реляционизм и холизм*, Деннет, говорит Рорти, сделал правильный шаг. Максима реляционизма выглядит следующим образом: «Если вы объяснили все реляционные свойства, имеющиеся у какой-либо вещи — все причины и все следствия, — значит, вы объяснили эту вещь»<sup>2</sup>. Вам больше не нужно искать ее сущность. Реляционизм деннетовской гетерофеноменологии предполагает, что все ментальные состояния должны определяться в терминах их каузальных отношений, а все, что выпадает из внешне фиксируемых реляционных связей, следует отнести к фиктивному. И еще. Когда найдено объяснение, *почему* люди говорят о внутренних свойствах опыта сознания, значит, найдено объяснение самих ментальных состояний. Например, в средние века некоторым людям приписывался «ведьминский инстинкт», который с развитием науки и медицины, использующих каузальные методы, был отнесен в разряд иллюзий; одновременно выявление социальных причин показало, почему эти иллюзии были в ходу.

Другая максима прагматизма — холизм (ее хорошо сформулировал Дэвид Дэвидсон и другие современные философы). Суть холизма состоит в отказе от проведения дилеммы между тем, чем объект является «сам по себе», и тем, что язык теории привносит в объект (или позволяет говорить о нем). Холизм провоз-

<sup>1</sup> Rorty R. Comments on Dennett // *Synthese*. V. 53. № 2. November 1982. P. 183. Далее: Rorty R. Comments on Dennett. P. ...

<sup>2</sup> Rorty R. *Holism, Intrinsicity, and the Ambition of Transcendence*. P. 185.

333

глашает единство «схемы — содержания» или теории и объекта. Главное отличие позиций Нагеля и Сёрля от позиций холистов состоит в том, что первые полагают, что объект — сознание — существует еще в виде событийности сознания, т.е. помимо наших объяснительных схем, вторые же это отрицают. Это различие можно сформулировать в виде дилеммы: «Должно ли наше объяснение охватывать отношения между мозгом, поведением и событиями в Картезианском Театре, или только отношения между мозгом, поведением и разговором о таких событиях?»<sup>1</sup>.

То, что Деннет называет «гетерофеноменологией» и объективным методом, Нагеля, конечно, никак не может устроить. В книге «Взгляд из ниоткуда» (1986) он отверг максимуму Витгенштейна «о чем нельзя сказать, не может мыслиться» и утверждает, что «источники философии являются до-вербальными и часто докультурными, и одна из ее наиболее трудных задач состоит в выражении в языке, без их потери, неоформленных, но интуитивно ощущаемых проблем». Отсюда

им делается вывод, что «содержание некоторых мыслей трансцендирует каждую форму, которую они могут принять в человеческом сознании»<sup>2</sup>. Понятно, почему во всей витгенштейновской линии современной философии ему видится «бедствие» — замыкание нашего рассуждения о мире «языковой тюрьмой». А склонность современных философов (Витгенштейна, Райла, Селларса, Дэвидсона, Деннета) к отказу от амбиции трансцендентности расценивается им как признак духовной деградации философии.

## 2. Деннет «сидит на заборе» между реализмом и ирреализмом

Определив Деннету место в лагере бихевиористов, реляционистов, номиналистов, инструменталистов, обобщенно обозначенное им «прагматизмом», Рорти с сожалением констатирует, что тот проявляет половинчатость в проведении прагматичес-

<sup>1</sup> Ibid. P. 186.

<sup>2</sup> Nagel Th. The View from Nowhere. Oxford, 1986. P. 11, 20, 102.

334

ких принципов. Истоки ее он видит в характерном для американской натуралистической традиции стремлении навести мосты между природой и культурой, реализмом и инструментализмом. Деннету, как и Пирсу, и Джемсу, и Дьюи, и многим другим американским философам дарвиновской традиции, говорит он, комфортно работается в случаях перелива и полутонов, без дихотомий и резких разграничительных линий. Однако это стремление к невозможному. Нельзя охватить одной схемой и природу и культуру: природа всегда предстоит перед нами в окультуренной нашими теориями форме.

Самый большой недостаток Деннета, определивший все остальные, Рорти видит в «мета-онтологическом восседании на заборе» между так называемым реализмом и так называемым ирреализмом. Это проявляется во многих вещах. Например, введя в свою гетерофеноменологию понятие «интенциональная установка», т.е. установку интерпретивизма, он встал на путь холизма, утверждающего единство «схемы — содержания». В то же время он сбивается с этого пути, пытаясь сохранить дистинкцию схемы и содержания, физического и интенционального, интенционального и реального. Например, он утверждает, что к исследованию сознания можно применять такой же эмпирический подход, какой применяют физики, причем предполагает, что это сделает исследование сознания теоретически уважаемым. И в то же время он проводит различие между исследованием электронов физиками и исследованием сферы интенционального гетерофеноменологами. В этих установках сказываются его сциентистские иллюзии и отступление от прагматизма<sup>1</sup>. Последовательный прагматист, твердо придерживающийся единства схемы и содержания, говорит Рорти, «не проводит различия между исследованием сознания и исследованием

<sup>1</sup> Рорти называет «сциентистами» людей, страдающих «научным шовинизмом»; следуя Карнапу и Попперу, они проводят демаркацию между просто разговором (политическим, моральным, литературным) и научным разговором. (Rorty and his Critics. P. 102).

335

Пруста или протонов. В исследовании любого объекта прагматисты видят одно и то же — способ справиться с вещью, соотнеся ее с чем-нибудь другим»<sup>1</sup>. Поэтому проблему сознания, как и другие проблемы, имеющие эмпирическую нагрузку, лучше отдать на откуп языковым играм ученых. Для философа «реальная проблема состоит не в том, как исследовать сознание эмпирически, а в том, чтобы убедить людей, что с методологической точки зрения она изучается точно так же, как любой другой объект»<sup>2</sup>.

Рорти нашел очень удачной метафору Деннета «центр нарративной гравитации» для объяснения того, почему в нашем языке фигурирует такое понятие, как «самость». К сожалению, ее антиэссенциалистский смысл подрывается сохранением различия между описаниями самостей и описаниями, используемыми в физике. «Мой принципиальный совет Деннету — ввести его утверждение, что самость есть «центр нарративной гравитации», в контекст более общего требования, что все объекты сходны с самостями, будучи центрами дескриптивной гравитации.

Нарративы как раз являются особой формой описания — описаниями, которые используют новеллисты и автобиографы, однако тип деятельности новеллистов принципиально не отличается от типа деятельности логиков, физиков и моралистов»<sup>3</sup>. Все эти люди плетут и заново переплетают веревки нарративов. Поэтому любой словарь для описания чего угодно — частиц или личностей — следует рассматривать как один среди многих других, полезный для одних целей и бесполезный для других. Приняв иную цель, нежели та, с которой первоначально создавался словарь, исходный словарь будет считаться фикцией. Словари Аристотеля, Лейбница или Бергсона, применявшиеся для разговора о движении, были вполне разумными и последовательными для своего времени, однако сегодня они не нужны. Поэтому

<sup>1</sup> Rorty R. Comments on Dennett. P. 186.

<sup>2</sup> Ibid. P. 183.

<sup>3</sup> Ibid. P. 189.

336

мы относим референты таких словарей к воображаемым сущностям<sup>1</sup>.

У Рорти вызывают скепсис и другие моменты концепции интенциональности Деннета. Выведя интенциональность за пределы «внутреннего Я» и отнеся ее к разряду естественных свойств, он поступил правильно. Ошибка состоит в сохранении дистинкции «реальное — интенциональное». Данное различие породило множество трудностей и не заслуживает того труда, который был потрачен на их устранение. Коли Деннет принял на вооружение холизм, предполагающий отказ от дистинкции «схема—содержание», тогда отпадает необходимость в использовании различия «интенциональное — реальное». Если согласиться с Д. Дэвидсоном, что истинность одних верований определяется другими верованиями, а не какими-то прочными фундаментами знания (а не согласиться с ним нельзя), тогда нет нужды называть какие-то объекты интенциональными и противопоставлять их реальным.

Отступление Деннета от холизма и прагматизма Рорти усматривает в его неоднозначных оценках репрезентирующих качеств повседневного разговора о ментальном или того, что называют фолк-психологией. Деннет полагает, что если словарь нашего обычного психологического языка (понятий фолк-психологии) помогает нам что-то понимать и предсказывать в межличностных отношениях, то происходит это, скорее, в силу его верной репрезентации реального, нежели чем просто кажущегося. Поэтому прагматические критерии предсказания и контроля для различения здравого и абсурдного недостаточны. В «Сознании объясненном» за несогласие с такой оценкой фолк-психологии он назвал меня, говорит Рорти, релятивистом, не признающим различия между фолк-психологией и астрологией, а мои прагматические взгляды отнес к «деревенскому верификационизму». Его собственный «городской верификационизм», считает он, является более софистичным в том отношении, что предполагает, что в каких-то случаях мы не

<sup>1</sup> Ibid. P. 186.

337

имеем, а в каких-то все же получаем «аккуратную репрезентацию», то есть проникаем За Занавес Явлений. В «Сознании объясненном» он пишет следующее: «Даже те люди, которые трансцендировали дистинкцию "схема—содержание" и увидели тщетность корреспондентных теорий истины (как, например, Рорти), должны согласиться с фактом, что в рамках естественного онтологического отношения мы иногда объясняем успех с помощью корреспонденции. Мы лучше проводим навигацию вокруг берегов Мейна, пользуясь современными морскими картами, нежели картой дорог Канзаса. Почему? Да потому, что первая карта точно представляет рискованные места, маркеры, глубины и линию берега острова Мейн, а последняя этого не делает. И теперь зададимся вопросом: почему мы лучше осуществляем навигацию среди опасностей интерперсональных отношений, используя фолк-психологию, а не астрологию?»<sup>1</sup>.

Приведенное Деннетом сравнение, замечает Рорти, легко опровергается старым прагматическим аргументом: мы не можем двигаться взад-вперед между нашими утверждениями об электронах и самими электронами, или приписыванием верований и самими верованиями, или сопоставлять их так, как мы сопоставляем кусочки карты и кусочки Мейна. Как сказал бы Витгенштейн, это



равносильно проверке сказанного в одной газете другой копией той же самой газеты. Деннет часто говорит, что термины «реалист» и «инструменталист» слишком расплывчаты, тем не менее он готов отнести себя в класс больших реалистов, нежели я, балансируя на мета-заборе, по одну сторону которого «ирреалисты», подобные (предположительно) мне, а по другую — «реалисты», подобные (предположительно) Дэвидсону. Я же хочу помочь убрать забор, отказавшись вообще использовать дистинкцию «реалист—ирреалист».

Свойственный Деннету прагматический настрой, сетует Рорти, должен был бы привести его к согласию со мной в

<sup>1</sup> Rorty R. Comments on Dennett. P. 195.

338

ненужности понятия «реальное». Но он на это не идет. Он сохраняет контраст между вещами, которые «реально реальны» и вещами, которые, как говорил Ройс, «не столь дьявольски реальны». Он продолжает держаться рейхенбаховского различия между статусом абстрактных понятий и идеальных объектов типа электрона (*illata and absrtacta*), говорит об «оснастке физического мира» и т.д. В отличие от «естественной онтологической позиции», которую отстаивает Деннет, моя прагматическая «естественная онтологическая позиция», говорит Рорти, не предполагает, что «реальность» в виде колесика входит в какой-либо механизм. В прагматизме вопрос о реализме формулируется просто: полезно ли вообще говорить о мире как пространственно локализованном, пространственно разделенном, материально ощутимом, наблюдаемом, легко идентифицируемом, состоящем из атомов, полезном для питания? К сожалению, говорит Рорти, Деннета не устраивает такое толкование реализма. «Я думаю, что "городской верификационизм" Деннета является немножко слишком городским. Он останавливается на полпути к цели, что, как мне представляется, является неуместной милостью к полуповерженному врагу»<sup>1</sup>.

Рорти призывает Деннета прекратить «восседать на заборе» и, расквитавшись с проблемой сознания, сделать следующий логически необходимый шаг и расквитаться с реализмом. «Ибо все приведенные им основания для освобождения от порабощающей и приносящей так много беспокойства картины сознания — "Картезинского Театра" — являются также основаниями для освобождения от порабощающей, приносящей только беспокойство картины человеческого исследования — "Проникновения за Занавес Явления"»<sup>2</sup>. И перестать тешить себя иллюзией, что в контрверзе «реализма—ирреализма» ему удалось найти третий путь, что-то вроде «умеренной-и-опосредующей разновидности реализма»,

<sup>1</sup> Ibid. P. 198.

<sup>2</sup> Ibidem.

339

отнеся мой вариант к «более умеренным-чем-умеренному ирреализму».

В половинчатости позиции Деннета, считает Рорти, не последнюю роль сыграл его излишний натурализм, настороженно относящийся к *историцизму*. На самом деле объяснить текст — значит, поставить его в определенный исторический контекст: соотнести его с ходячими конвенциями, происходящими событиями, другими текстами, ситуациями собственной жизни и многим другим. Ничего такого нет у Деннета. Во многом это объясняется его нежеланием входить в *метафилософские* сферы. Отстраненность от них помешала ему увидеть сходство его наступления на аристотелизм с таким же наступлением в рамках европейской традиции герменевтики. Последняя кажется ему «опасной территорией», куда лучше не входить. Если бы он рискнул туда войти, он нашел бы для себя много полезного и увидел, что в деятельности Хайдеггера, Деррида, Гадамера и других нет ничего опасного, а по сути они продолжают операционистскую и антиэссенциалистскую традицию, которую заложил Мах, продолжили Дьюи и Гудмен и к которой на законном основании мог бы примкнуть и Деннет.

## Деннет о Рорти

### 3. О «привилегированных» репрезентациях старого когнитивизма и «непривилегированных» репрезентациях нового когнитивизма

Реакцию Деннета на комментарии Рорти в целом можно охарактеризовать как протест против его слишком прямолинейных и общих суждений. «Я вношу коррективы в сказанное Рорти и пытаюсь накинуть мокрое одеяло на его слишком провокативные интерпретации, хотя не исключаю, что поступаю так по причине того, что остался близоруким рортиев-

340

ским революционером с макаронобразными ногами, при всей его энергичной апелляции ко мне и стремлении извлечь мораль из последней моей работы»<sup>1</sup>.

Как уже говорилось, Деннет в целом высоко оценивает работу Рорти в области метафилософии и, в частности, его обзорную работу «Contemporary Philosophy of Mind» (1982). Вместе с тем, замечает он, тактика широких мазков, будучи сама по себе полезной и неизбежной в истории философии, чревата опасностью упрощения ситуации, если смазывать тонкости в позициях рассматриваемых авторов, не замечать их замешательства перед трудностями проблем, удерживающих от категоричных суждений. К сожалению, Рорти злоупотребляет этой тактикой и часто «перебарщивает». Представленные им размышления, говорит он, подтолкнули меня провозгласить полезный герменевтический принцип или «*Rorty Factor*»: возьми все, что говорит Рорти о чьих-то взглядах, и умножь это на 0,742. «Подобно многим другим революционерам до него, Рорти оказался в затруднении: то ли следует провозгласить победу, т.е. объявить, что победа неотвратима, то ли умолять вас присоединиться к трудной и неясной борьбе против сил тьмы. Я спрашиваю себя: являюсь ли я *номиналистом*? декларирую ли я смерть теорий сознания? считаю ли я себя Деревенским верификационистом или мне надлежит стать им в конце концов? На все эти вопросы у меня всегда возникает желание ответить — *не совсем*»<sup>1</sup>.

Будучи не только философом, но и когнитивным ученым, Деннет хорошо осведомлен о трудностях исследования загадок сознания. Над всяким элиминативизмом, и над элиминативизмом Рорти в том числе, черной тучей висит асимметрия между «перволичностным» и «третьеличностным» отношением к миру. Массу вопросов вызывают квалиа, пространственно-временная локализация актов сознания, визуальное воспри-

<sup>1</sup> Dennett D. Comments on Rorty // *Synthese*. V. 53. № 2. Nov. 1982. P. 350.

<sup>2</sup> Ibid. P. 349.

341

ятие, не говоря о переходе довербального уровня на вербальный. Философу, предупреждает он, нужно быть предельно осторожным в обобщениях.

Конечно, Рорти прав, что идеи о привилегированном доступе к своему собственному сознанию, о достоверности внутренних репрезентаций, о зеркальности отражения — это мертвый груз прежней философии. Тем не менее никогда не следует забывать, что наши когнитивные способности являются продолжением способностей природы, и нелепо думать, что они устроены для обмана. Прагматический или «деревенский верификационизм» Рорти слишком просто рисует ситуацию познания. В действительности назрело время объяснить на новом уровне знания фактическую репрезентационную работу сознания.

Когда Рорти утверждает, что у знания не существует никакого твердого эпистемологического основания, он только отстаивает особый случай антифундаменталистской доктрины, которую развивали Куайн, Витгенштейн и другие авторы. Однако еще нужно доказать, удалось ли кому-либо стать последовательным витгенштейнцем или куайнцем в особом случае знания своего собственного сознания. Декларировать банкротство *привилегированных репрезентаций* своего собственного сознания, как это предлагает Рорти, еще рано, потому что все мы сохраняем для себя некоторую авторитетную привилегию в отношении собственных взглядов. В частности, у нас на руках нет еще детального и удовлетворительного диагноза неустраняемой асимметрии

отчетов от первого и третьего лица. А тот, который имеется, все еще противоречив и на горизонте революционных новаций Рорти висит черной тучей, которую он недооценивает. Поэтому его критический аргумент против Нагеля и Сёрля на основе аналогии между сознанием и движением не совсем убеждает.

Недооценку асимметрии Деннет связывает, в частности, с отказом Рорти различать два совершенно разных смысла *интуиции*. С одной стороны, некоторые философы говорят об *интроспективных* интуициях, как о направленных вовнутрь

342

себя кусочках до-лингвистического, достоверного знания. Доказать такие «правильные» интуиции вряд ли возможно. С другой стороны, существуют интуиции в смысле «до-теоретических интуиций», или «грамматических интуиций», или интуиций, опираясь на которые мы воспринимаем некоторые заявления как «контринтуитивные». Они тоже уязвимы, однако изначально игнорировать их нельзя. Интуиции, о которых говорит Сёрль, сами по себе не являются интроспективными интуициями.

Деннет отнюдь не поддерживает перехлесты Рорти в оценках Нагеля и Сёрля и сомневается, что их, «как неандертальцев каменного века», можно с легкостью устранить с пути. Он напоминает, что его герой У. Селларс, известный своими разоблачениями «мифа данного», испытывал замешательство перед квалиа — сырыми чувствами или качественными состояниями. Что его внимание было приковано к абсолютно *другому* виду интуиции, с тем чтобы предложить фундаментальную ревизию всех наук — не только психологии и биологии, но и всей физики, — для того чтобы справиться с *квалиа*. Отрицая интуиции в первом смысле, он все же чувствовал, «что что-то подлинное упускается в том взгляде на сознание, который исповедуем и Рорти, и я сам»<sup>1</sup>.

Немаловажно и такое соображение. Если совсем отказаться от философского исследования репрезентаций, тогда всю работу по анализу приобретения знания мы должны передать науке. Однако стоит обратиться к ученым, изучающим физическую сторону сознательных процессов, и мы обнаружим у них не меньшее замешательство. Попросите их рассказать, каким образом любая живая или неживая вещь имеет диспозиции произносить что-то, познавать что-то, действовать на что-то, и окажется, что их объяснения сводятся к тому, что такая вещь либо сама должна быть системой *репрезентаций*, либо содержать ее в себе. То есть они снова выходят на философскую проблему репрезентаций.

<sup>1</sup> Ibid. P. 350.

343

Удалось ли нам сегодня оставить позади себя все прежние ужасы репрезентационных теорий сознания? Не разбили ли мы зеркало природы только для того, чтобы обнаружить за ним другое зеркало? На эти вопросы Деннет отвечает отрицательно. На самом деле мы нуждаемся в новом понятии *когнитивизм* и новом смысле *репрезентивизма*.

Ключевое значение в новом когнитивизме играет рассмотрение сознательных состояний как результата интерпретации отношений между поведением, окружающими условиями и диспозициями (включая эволюционные). Новый когнитивизм — это эволюционный когнитивизм. То есть они являются состояниями всей системы личности — поведения—мозга—окружающих условий, при том ограничении, что система, имеющая цели и верования, является достаточно рациональной в ее действиях по достижению целей.

Новый когнитивизм не отвергает аргументы Гудмена против фундаментализма и не считает репрезентации конечными, не подверженными корректировке основаниями, или источниками всех значений и интенциональности. Хотя репрезентации эксплицитно и буквально имеют место *в головах* тех, кто имеет верования, интенции и пользуется языком, однако им можно приписать почти такой же статус, который приписывается таким переносчикам знания, как книги, карты и другие репрезентивистские *протезы* нашей культуры. Подобно тому как семантическая или герменевтическая интерпретация прикрепленных к доске кусочков текста зависит от нашего

приписывания им интенциональности, семантическая интерпретация кусочка «ментальной» репрезентации в мозге — надписи на «языке мысли» — зависит от такой же семьи интенциональных атрибутов. Точно так же, как мы прибегаем к книгам и словам для коммуникации друг с другом, мы используем собственные мозги для коммуникации с самими собой — с нашими прошлыми, настоящими и будущими самостями.

Кому-то идея внутренних репрезентаций, имеющих только выводимую, прагматическую интенциональность или семантическую, говорит Деннет, может показаться слабостью. Но это

344

только на первый взгляд. Например, считается, что компьютер — не что иное, как автоматическая книга, имеющая только те смыслы, которые вложили в нее интерпретаторы. Заключение в этом аргументе вывод состоит в том, что только мы — неинтерпретированные интерпретаторы — являемся источниками появления всех реальных смыслов. В противоположность такой расхожей линии рассуждения новый когнитивизм обращает внимание на следующее: не существует ни неинтерпретированных интерпретаторов, ни привилегированных репрезентаторов. Вывод, следующий из этой точки зрения, выражен в простом лозунге, что наши мозги являются органическими компьютерами и включены в информационные процессы природы. Поэтому можно сказать, что «состояния мозга *точно так же, как компьютеры*, могут быть интерпретированы внешними наблюдателями с помощью некоего рода герменевтической процедуры, как имеющие содержание... Мы одновременно и творцы и творения такой интерпретации, и ничего другого не существует за пределами этой активности»<sup>1</sup>. В «Сознании объясненном» он такими словами рисует нашу включенность в информационные процессы природы: «Когда в моменты глубокого погружения в самосознание мы восхищаемся великолепным богатством нашего сознательного опыта, то, чем мы восхищаемся, в действительности является богатством мира вне нас во всех его приводящих в восторг деталях. Мир не «входит» в наше осознающее себя сознание, но просто является доступным»<sup>2</sup>.

Какие бы изощренные аргументы ни приводили философы в доказательство отсутствия фундаментов у знания о теоретической нагруженности данных, о первичности теоретического по отношению к эмпирическому и отсюда о невозможности реализма, упрямые факты говорят, что в одних случаях эти аргументы работают, а в других нет. Ссылки Рорти на авторитет Витгенштейна, Гудмена, Куна и других антифундамен-

<sup>1</sup> Ibid. P. 355.

<sup>2</sup> Dennett D. *Consciousness Explained*. P. 407—408.

345

талистов, строго говоря, не в состоянии поколебать очевидность этих фактов.

Не все так просто и с прагматическими критериями, к которым апеллирует Рорти. Конечно, говорит Деннет, моим важнейшим мотивом введения в эпистемологический арсенал «интенциональной установки» является ее прагматическая объяснительная возможность: эффективность, предсказательность, установление контроля и др. Однако толковать «интенциональную установку» только в социологическом ключе (только как интерпретивизм) неверно. Не следует забывать, что стремление к прагматической эффективности присуще не только социальному человеку. Сходные с интенциональной установкой признаки можно наблюдать и в природе: хищники интерпретируют поведение других животных прагматически и интенционально, приписывая им рациональные верования и желания, в противном случае охота была бы невозможна. Поэтому сведение интенциональной установки только к инструментальности — социальному контролю и верификации верований на основе прагматической эффективности, как это делает Рорти, существенно упрощает картину. Одним словом, в случае рационального поведения и человека и животного взаимодействие факторов столь перекручено, что объяснить его с помощью предлагаемого Рорти прагматического критерия вряд ли возможно. Деннет называет его упрощенным «деревенским верификационизмом».

Многих людей, говорит Деннет, воинствующий инструментализм Рорти приводит в содрогание, особенно пугающе звучат его заявления, что наше представление о себе как о личностях — это

психологически приятный самообраз, ничего общего не имеющий с истиной. Рорти поощряет этот страх провокационными заявлениями о социологических причинах такого самообраза. Конечно, знание социологических причин возникновения такого самообраза, говорит Деннет, важно, но дело в том, что такое представление *также истинно*. Иначе говоря, это такой же истинный образ, как и сформировавшийся в ходе истории образ Земли в виде вращающейся вокруг Солнца планеты. «Если человек может принять все эти «открытия», а затем

346

*в принципе* релятивизирует такие взгляды соотнositельно с «интересами» и «языковыми играми» — деятельность, в отношении которой, несмотря на риторику Рорти, Патнэма, Куна и Гадамера, я остаюсь агностиком, — тогда, по меньшей мере, давайте никогда не будем забывать напоминать самим себе о том, что наше видение самих себя как людей познающих и заботящихся в эпистемологическом смысле не хуже, чем наше видение воды, как состоящей из H<sub>2</sub>O»<sup>1</sup>.

#### 4. Метафоризм, базисный инстинкт и «вегетарианская» концепция истины

Как мы уже говорили, и Деннет и Рорти — мастера по изобретению метафор, однако они по-разному оценивают их эпистемические возможности. Согласно Рорти, человек по своей натуре — поэт, осваивающий окружающий мир с помощью различных метафорических языков. Наука — не исключение: это один из подвидов метафорического «переописания» действительности, ее словарь не в большей мере обладает привилегией «доступа к реальности», нежели словарь искусства. Тезис Гудмена о невозможности обоснования знания фундаментами — доказанная истина. А поскольку ни у науки, ни у философии нет под ногами твердой почвы, никакая теория не репрезентирует «Реальность» и не открывает «Истину». Деннет, думая, что он возвышает статус философии, уподобляя ее науке и обособляя ее от литературы, вводит себя в заблуждение: движимый цеховыми интересами самоутверждения имиджа вида философии, с которым он работает, он просто изобретает новые метафоры. Но никак не желает это признавать, и отсюда все его непоследовательности.

Одобрение и похвалу Рорти вызвала концовка книги «Сознание объясненное», где тот пишет следующее:

«Я не заменил одну метафорическую теорию — Картезианский Театр — другой, неметафорической («буквальной, на-

<sup>1</sup> Dennett D. Comments on Rorty. P. 355.

347

учной») теорией. В действительности все, что я сделал, — это поставил на место одного семейства метафор и образов другое семейство, поменяв Театр, Наблюдателя, Центральный Источник значений, Фикцию на Мягкую Программу, Эффективные Машины, Многократные Наброски, Пандемониум Гомункулусов. Вы скажете, что это только война метафор, однако метафоры не есть «только» метафоры, метафоры являются инструментами мысли. Без них невозможно мыслить о сознании, поэтому очень важно вооружить себя лучшей из имеющихся в нашем распоряжении системой инструментов. Оглянитесь вокруг и посмотрите, что мы создали с помощью наших инструментов. Можете ли вы все это вообразить себе без них?»<sup>1</sup>

Данный пассаж, считает Рорти, можно считать трезвой самооценкой инструменталиста: Деннет подавил в себе естественное стремление объявить, что он, наконец, точно обрисовал внутреннюю сущность сознания. Вместо этого он говорит, что все, что он сделал, — это «заменял одну семью метафор и образов на другую». «Мне хотелось бы, чтобы он предпринял еще один шаг в этом направлении и добавил, что такие инструменты — это все, что когда-либо может дать нам исследование»<sup>2</sup>.

Однако Деннет категорически отказывается его предпринимать. Он против использования какого-либо положения его теории, в частности его толкование метафор, в качестве «взрывоопасного пункта» в проектах разрушения представления о возможности достижения объективности. «Я



никогда не предприму этот шаг; хотя метафоры действительно являются незаменимыми инструментами мысли, однако их нельзя считать единственными. И микроскопы, и математика, и сканнеры — такие же инструменты. Конечно, всякое исследование пред-

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 455. Нужно сказать, что у ряда комментаторов такая концовка книги вызвала недоумение: как такое может говорить натуралист?

<sup>2</sup> Rorty R. *Holism, Intrinsicity and the Ambition of Transcendence*. P. 198.

348

ставляет собой приобретение нами чего-то, к чему мы стремимся, но стремимся мы — иметь истину о чем-то значимом для нас»<sup>1</sup>.

Однажды, говорит Деннет, он совершил ошибку, согласившись, чтобы его включили в компанию инструменталистов. «Больше я никогда не буду ее допускать»<sup>2</sup>. Ярлык «инструментализм» используется в разных значениях и часто вносит путаницу в понимание познания. По традиции инструментализм считается радикальной доктриной, помещающей в одну лодку и электроны, и китов, и экватор. Однако есть вполне разумный его смысл, которым пользуются нормально мыслящие ученые — инструментализм в отношении центров гравитации, параллелограммов сил и т.п. И еще. «Я не убежден, что истинное относительно сознания является истинным в отношении всего другого. Рорти считает мою атаку на Картезианский Театр мелкой перестрелкой в широкой атаке на "порабощающую, но приносящую только нескончаемые трудности картину человеческого исследования: Проникновения за Занавес Явления". Я с этим не согласен»<sup>3</sup>.

У российского читателя полемика Деннета и Рорти по поводу метафоризма может вызвать естественный вопрос: а зачем нам вообще нужно менять метафоры, с помощью которых мы более или менее сносно ориентируемся в собственных мыслях и ощущениях и межличностных отношениях? Ради чего стоит отказываться от старых, удобных и хорошо работающих менталистских метафор Картезианского Театра ради неудобных метафор «Модель Множественных набросков»? Какие причины, кроме психологической усталости от старых метафор и желания чего-то новенького, толкают к смене словаря? В научных дисциплинах новые метафоры обычно

<sup>1</sup> Деннет Д. Почему нам важно понимать это правильно. Постмодернизм и истина. С. 484. Сноски даны на перевод в Приложении 1 настоящего издания.

<sup>2</sup> Dennett D. *Back from Drawing Board // Dennett and His Critics*. P. 210.

<sup>3</sup> *Ibid.* P. 234.

349

вводятся при обнаружении какого-нибудь нового явления. Например, компьютерщики срочно позаимствовали у микробиологов термины «вирус» и «эпидемия» для обозначения распространения по сетям программного сбоя. Какие вообще резоны подтолкнули Деннета и Рорти разрушать старый декартовский образ?

На эти вопросы Деннет и Рорти отвечают по-разному. Хотелось бы обратить внимание на эту разницу, поскольку за ней скрывается различие *идеологических установок* этих двух мыслителей. Рорти полагает, что метафоры служат для ориентации и адаптации в меняющейся жизни и поэтому сами тоже должны меняться. Смена метафор, говорит он, нужна «в силу обычных куновских причин». Мы испытываем потребность в новых интуициях, поскольку старые никуда нас не привели, и нам нужен новый язык, поскольку идея репрезентаций породила огромное количество не поддающихся решению контроверз и фантазий, висящих тяжелым грузом на теле философии. Хотя, замечает он, в конечном счете все зависит от нашего собственного выбора «куда мы хотим идти» — хотим ли и впредь лелеять трансцендентность или хотим навсегда отказаться в ней искать архимедову точку опоры и обратиться к ориентирам, которые подсказывают жизнь и практика нашего сообщества<sup>1</sup>.

То есть, по его мысли, к смене метафор нас толкает не столько даже прагматическая необходимость для обозначения нового явления, сколько психологическое *желание обновления*.

У Деннета другая позиция. Метафоры, конечно, являются инструментами, и далеко не всегда они помогают нам проникать «За Занавес Явления»; часто они заводят в тупики. В то же время

хорошие, удачные метафоры проясняют проблемную ситуацию и помогают ее разрешить и, что особенно ценно, могут становиться инструментами предсказания и открытия нового. Меняя один словесный инструмент на другой, мы

<sup>1</sup> Rorty R. *Holism, Intrinsicity, and the Ambition of Transcendence*. P. 199.

350

приобретаем все более точное знание. Рорти, конечно, прав, говоря, что нами всегда движет желание. Однако не следует забывать, что *первичным* желанием человека всегда было *желание иметь Истину*. Желание постичь истину отнюдь не произвольно: оно является нашей *биологической* потребностью. Для любого организма нет ничего более существенного, нежели иметь адекватную информацию об окружающей среде и о самом себе. Иначе он погибнет.

Частенько забывают, говорит Деннет, что, споря об истине и ее критериях, философы спорят о способах недопущения инфляции истины об истине в Истину об Истине, разбухания ее в некоторую абсолютистскую доктрину, накладывающую непререкаемые требования к нашим системам мысли. То же самое можно сказать относительно дебатов о реальности времени или реальности прошлого. Какими бы софистичными ни были разговоры (и разногласия) относительно возможности логического доказательства реальности прошлого, они не подрывают такие, например, утверждения: «Джек Руби выстрелил и убил Ли Харви Освальда в 11.25 по далласскому времени 22 ноября 1963 года».

Когда я припер Рорти к стенке по вопросу истины, говорит Деннет, тот согласился признать наличие полезного понятия истины, остающегося незыблемым после всех вездыхальных философских опровержений, например, при решении вопроса о виновности или невинности лица в совершении преступления. Признав это, Рорти тем самым признал пропасть — и ее важность — между явлением и реальностью, между театральными словесными упражнениями, развлекающими без претензии на сообщение истины, и изысканиями, нацеленными на истину и часто достигающими ее. Правда, последнюю он все же низвел до разряда скромного «вегетарианского» понятия истины. Однако для ученых, никогда не стремящихся доводить все до абсурда, она чрезвычайно важна, и они всегда работают с ней. И это не случайно.

У скромного «вегетарианского» понятия истины на самом деле, говорит Деннет, имеется мощный фундамент — *базис-*

351

*ный инстинкт* или дарвиновское правило всего живого — делай все правильно и не допускай ошибок. Миллиарды организмов на нашей планете в течение более чем трех миллионов лет вовлечены в игру: спрячься и ищи. Для них это не только игра. Это дело жизни и смерти. Для организмов нет ничего более существенного, нежели иметь адекватную информацию об окружающем мире, но они, как правило, делают это бессознательно. В случае сбоя инструментария или ложного восприятия ситуации у них нет ресурсов для эффективного исключения ошибки и гибели. *Распознавание* различия между явлением и реальностью — человеческое достижение. Только человек наделен сомнением и эпистемическим зудом к поиску лекарства от него — лучших методов достижения истины. Земля Лжецов существует только в загадках философов: традиция системы Ложного Календаря для неверной записи хода времени не сложилась ни у кого. «Принцессой, правящей балом», на котором присутствуют ложь, обман, надувательство, самообман, продолжает оставаться Истина; без нее все другие танцоры теряют смысл<sup>1</sup>.

## 5. Регулятивный идеал философии: объективность или солидарность?

Рассказывая о полемике Деннета и Рорти, мы не будем производить теоретическую разборку и оценивать, кто из них прав, а кто не прав. Поскольку наша работа дескриптивная, содержащая много цитат, читатель располагает возможностью самому произвести анализ и делать выводы. К нарисованной нами картине мы хотели бы добавить еще пару штрихов, касающихся корневой оппозиции.

Как можно было видеть из представленного материала, разногласия Деннета с Рорти выходят за

рамки эпистемологии и завязаны на разном понимании регулятивных идеалов

<sup>1</sup> Деннет Д. Почему нам важно понимать это правильно. Постмодернизм и истина. С. 487.

352

философии. Для Рорти принципиальным является тезис, что все философские проблемы — отношения сознания и тела, субъекта и объекта, реального—ирреального, свободы и детерминизма; то есть, вопросы, которыми занимается Деннет, не являются подлинными, естественными или, как он говорит, «преднайдеными». Возникнув в результате языковой путаницы и «языковых игр», они незаконно обрели статус теоретических.

Деннет категорически не соглашается с этим тезисом. При всем уважении к Витгенштейну, ему чужд образ философии как «лингвистической игры». Действительно, некоторые проблемы возникли в результате языковой путаницы и подлежат элиминации. Тем не менее очень многие являются подлинными, естественными и решаемыми. Подлинность философских проблем удостоверяется фактом их постоянного восстановления и «открытия» в процессе движения любого знания и их вхождения в фундамент знания. Часть из них переходит в сферу науки. Тем не менее роль философии не сводится к «повивальной бабке», в ней существует свой собственный прогресс: плохо сформулированные вопросы становятся лучше сформулированными, старые проблемы обретают новые смыслы, обрастают семейством подпроблем и т.д. Перекочевав в «тело» науки, философия продолжает работать и там. Ибо все вопросы ее концептуального основания относятся к ведомству философии. Этим объясняется эвристичность философских гипотез. Сомнение в эвристичности философских гипотез негативно сказывается на знании. Например, сегодня прогресс в когнитивных науках, в теориях искусственного интеллекта, в нейробиологии тормозится из-за отсутствия в этих дисциплинах перспективных философских идей и серьезной философской критики их концептуальных оснований. Следует заметить, что у Рорти и Деннета примерно одинаковый замысел: преодолеть идущий от Декарта раскол мира на *res extensa* и *res cogitans*, а культуры — на науки о природе и науки о «духе», и доказать возможность единого поля для обсуждения разных проблем. Однако векторы его реализации

353

направлены в разные стороны. Деннет, взяв курс на теоретизацию философии, стремится обосновать непрерывность мира и единство знания ценой отказа от жесткой дилеммы лингвистического и нелингвистического знания, натурализма и социологизма. Рорти взял курс на «историцизм», языковый релятивизм и «философию-как-литературу». Различие между «мягкой» гуманитаристикой и «твердыми» науками — естествознанием и математикой, с его точки зрения, пропадает, если посмотреть на них как на различные типы словарей, изобретенные для меняющихся прагматических целей. Кажущаяся пропасть между «сциентистской» аналитической философией и «литературно» ориентированным постмодернизмом тоже исчезает, если посмотреть на их «серьезную» и «фривольную» работу как на литературное творчество; в этом случае различие сведется к стилевым и языковым особенностям<sup>1</sup>.

Подводя предварительный итог полемике между Деннетом и Рорти, нельзя не признать, что на первый взгляд позиция Рорти выглядит достаточно прочной в сравнении с «сидящим на заборе» Деннетом. Работая в каркасе лингвистически-прагматистской (историцистской) парадигмы, Рорти строго следует ее правилам. (По сути, принял тезис Р. Карнапа, что в рамках избранной теории можно ставить только внутренние для нее вопросы, внешние вопросы будут к ней иррелевантны.)

<sup>1</sup> Правда, в последнее время Рорти признал и другое отличие — и оно существенно. «Нет сомнения,— говорит он, — что имеется смысл, в котором аналитическая философия действительно более научна, нежели другие типы философии. Большинство аналитических философов являются разгадывателями загадок (*puzzle-solvers*) в том смысле, в каком Кун говорил, что естественные науки были *puzzle-solvers*. Они обнаруживают противоречия в наших интуициях и ищут способы их разрешения точно таким же образом, как ученые-естественники находят противоречия между теориями и наблюдениями, а затем обдумывают способы разрешения этих противоречий. Философы неаналитического склада, как правило, не занимаются решением загадок» (*Rorty and his Critics*. P. 107).

354

Вопрос о реализме для него стоит в виде дилеммы: либо следовать максиме Витгенштейна «о чем нельзя говорить, о том нельзя и мыслить», либо постулировать трансцендентное и способность проникать За Занавес Явлений. Центризм и компромиссы для него неприемлемы: философия сознания может быть представлена либо монизмом, либо дуализмом с его трансцендентностью. Нельзя поступать так, как это делает Деннет: придерживаться интерпретивизма, холизма и реляционизма и одновременно допускать какие-то элементы традиционного реализма.

Поэтому он поддерживает только те элементы в творчестве Деннета, которые связывают его с традицией Витгенштейна— Райла—Гудмена, и прежде всего — «зомбизм» в подходе к сознанию и интерпретивизм в понимании объяснения. И отказывается принять всерьез (оценивает как «путаницу» или «сидение на заборе») все новации Деннета, выходящие за рамки этой традиции: идею интенциональности, различие физического и интенционального и, конечно, идею Матери-Природы как первичного дизайнера языка, артефактов и культуры. Иначе говоря, те элементы, с помощью которых тот пытается состыковать социолингвистическую модель с натуралистической. (Еще меньше устраивают Рорти биологистские модели, например, Н. Хомского и Дж. Фодора. Этим авторов, по его словам, больше интересуется не то, что связывает нас с Платоном, а то, что связывает с шимпанзе. Защищая историцизм и холизм схемы/ содержания, он называет тщетными все попытки «выпрыгнуть» из культуры и «прорваться» к самой природе.)

Тем не менее, при всей ее уязвимости, позиция Деннета нам кажется перспективнее. Деннет относился с глубочайшим уважением и к Витгенштейну и Райлу, к их идее, что многие философские проблемы возникли из-за путаницы в языке, тем не менее он увидел в использовании менталистских идиом естественного языка поддержку не для грубого бихевиоризма, но «эволюционного когнитивизма»; что для понимания поведения мы должны прибегнуть к более широкому репертуару ментальных состояний и процессов и что последние лучше

355

всего понимать в рамках эволюционизма и естественного отбора. Обращение к эволюционизму означало кардинальный теоретический сдвиг в философии естественного языка.

Поэтому его не устраивает стратегия в рамках только социолингвистической парадигмы, что свойственно Витгенштейну и Рорти, ибо в этом случае элиминируется сознание и природная компонента человеческого бытия превращается в чистый концепт или «вещь в себе». Его информационно-процессуальная теория сознания и эволюционный когнитивизм имеют целью доказать, что человек с его лингвистически оформленным сознанием и социальной организацией включен в информационную систему природы и ее протоинтенциональность. Доказать это очень трудно, в распоряжении современного человека нет еще достаточного количества фактов, на которые можно твердо опираться. Тем не менее философские гипотезы о соединении биологического и социального нам кажутся более обнадеживающими, нежели чисто социологические гипотезы.

## 6. О вирусе постмодернизма и «западной» науке

В 1998 году на XIX Всемирном философском конгрессе (Бостон, 13 августа 1998 г.) Деннет выступил с докладом «Почему нам важно понимать это правильно. Постмодернизм и истина» (текст доклада дан в Приложении 2 этой книги). Любопытный читатель, который проявит интерес к изменению тональности прочитанных в последние два десятилетия заглавных для всемирных философских конгрессов докладов, непременно заметит в них нарастающий идеологический мотив — о важности идеалов Просвещения, рационализма, истины и философии как мудрости. (Если он поинтересуется тональностью докладов более раннего периода, то заметит скорее антиидеологизм: по понятным причинам мировое философское сообщество с опасением наблюдало распространение по миру идеологизированной марксистской философии.) Риторика верности ценностям Просвещения присутствует и в

356

докладе Деннета, причем, может быть, более открыто, нежели в докладах такого же уровня на предшествующих конгрессах, поскольку ясно обозначился противник. Речь в нем идет о нависшей

над философией большой опасности — распространении по миру постмодернизма. Постмодернизм сравнивается с «болезнетворным вирусом», реплицирующей себя вредной идеей или «меме». Особую опасность этот вирус представляет для не имеющих к нему рационалистического иммунитета стран третьего мира.

Деннет начинает доклад с придуманной им истории о том, как группа американских исследователей без особого злого умысла занесла вредоносный вирус в страну третьего мира, нанеся огромный урон здоровью людей и ухудшив политическую ситуацию: была подорвана единственная демократическая сила страны и укрепилась позиция традиционного деспота. Эти исследователи не были биологами, агрохимиками или медиками, они были постмодернистами. Под флагом мультикультурализма и релятивизма они санкционировали местные традиции, право любого сообщества на свою «Истину», какой бы она ни была архаичной, и сеяли семена предубеждения против всякого универсализма и прежде всего — против «западной» науки, рационализма и Истины. Мораль этого рассказа Деннета — умышленный или неумышленный подрыв сил «западной» науки и рационализма в странах третьего мира — оборачивается для живущих там людей бедствиями.

Во второй, уже не придуманной истории Деннет рассказывает о том, как однажды, когда он был еще молодым профессором, к нему обратился известный литературный теоретик с просьбой снабдить его «какой-нибудь эпистемологией», лучше неординарной. По-видимому, в том сезоне в кругу *literaty* было модно щеголять какой-нибудь «эдакой» эпистемологией. Мораль этого рассказа — в среде *literaty* нет ощущения грани, отделяющей серьезную теоретическую работу от эстетически экстравагантных языковых игр. Самое печальное состоит в том, что к такому умонастроению тяготеют и некоторые философы. Многие люди внутри и вне философии не видят,

а  
357

некоторые активно отрицают образовавшуюся в философии *пропасть* между серьезностью и фривольностью. В итоге философия превращается в литературную игру, искусство, а философские высказывания — просто в мнения, с легкостью заменяемые в другой игре на иные мнения.

Будь постмодернистская мода поветрием в узком кругу интеллектуалов, не было бы ничего страшного. Но постмодернисты выходят за пределы академического мира и стремятся влиять на умы широкой публики. Поэтому уместно, считает Деннет, поставить вопрос о применении к философии принципа, принятого в прикладных дисциплинах, в частности в медицине, — «не навреди». Нам нужно принять позиции и правила, действующие в прикладных дисциплинах, возлагающие ответственность на нас за наши слова и требующие гарантии их безопасности<sup>1</sup>.

Деннет делает еще более резкое движение: с трибуны Конгресса он заявил о своем идейном разрыве с «другом Диком» — Ричардом Рорти, дрейфующим в сторону постмодернизма. Как говорится, «Платон, ты мне друг, но истина дороже». «К моему глубокому сожалению, среди людей, вносящих лепту в распространение этой позиции, находится мой

<sup>1</sup> Деннет Д. Почему нам важно понимать это правильно... С. 481. Российскому читателю, помнящему историю использования советской системой марксистской философии в качестве ее «идеологического оружия» и последующего низвержения марксизма новой системой власти, такое требование может показаться наивностью американского профессора. Но это только на первый взгляд. Марксизм официально был низложен, однако при этом не было всерьез принято, т.е. не стало академической нормой, правило «не навреди». В нашей стране, как и в странах третьего мира, не имеющих сильного рационалистического иммунитета, гостеприимно были распахнуты двери (и умы не очень сведущих людей) хлынувшим в большом количестве с Востока и Запада интеллектуальным товарам, несоместимым с идеей истины и образования как Просвещения, не говоря уже об откровенно шаманских товарах.

358

друг Дик Рорти. Более четверти века между Ричардом Рорти и мною велась конструктивная полемика. Рорти открыл нам новые горизонты современной философии, пронизательно показав нам, философам, множество вещей, касающихся истоков наших проектов в философских проектах далекого и близкого нам времени, и в то же время он жестко описывает и предписывает будущие



шаги, которые нам следует предпринять. Есть один пункт, относительно которого между нами нет никакого согласия — пока нет согласия, — и он связан с многолетним стремлением Рорти доказать, что дебаты философов об Истине и Реальности в действительности ликвидируют пропасть и фактически дают лицензию на скатывание в некоторую форму релятивизма. В итоге Рорти сообщает нам, что все эти дебаты не что иное, как "разговоры", и выбор в них той или иной роли определяется только политическими, или историческими, или эстетическими основаниями<sup>1</sup>. В целом такая риторика и такой тон несвойственны стилистике Деннета. Его, как мы уже не раз говорили, интересуют философские «пазлы», а не метафилософские *разговоры* о философии, тем более о ее идеологии. И только очень веские основания могли подвинуть его на такой разговор. Пространство идеологического несогласия четко обозначилось тогда, когда стало ясно, что вектор Рорти на метафоризацию, эстетизацию и релятивизацию философии способствует распространению по миру антисциентистских настроений.

В многочисленных обличениях науки, говорит Деннет, подчас забывают говорить о существовании организованного энциклопедического собрания огромного количества согласованных научных фактов, без которых немыслима цивилизованная жизнь. Но дело даже не в накопленных фактах и впечатляющих прикладных результатах науки. Главное состоит в том, что ей, как ни одной другой форме культуры, присущи такие уникальные черты, как самокритичность, самофальсифицируемость, пересмотр собственных посылок и

<sup>1</sup> Там же. С. 483.

359

оснований. Такими чертами не обладает никакая другая форма культуры. И в силу этого она является наиболее объективной и прогрессивной из имеющихся технологий достижения Истины. Конечно, Рорти прав, говоря, что не существует архимедовой точки опоры, позволяющей благословлять науку, что оценщики сами находятся внутри научного сообщества и пользуются принятыми в нем нормами. Однако он глубоко заблуждается, относя сформировавшиеся в западной цивилизации стандарты научной рациональности к «местечковым». Наука, как, может быть, ни один другой изобретенный человеком инструмент для получения информации о мире, включена в процесс естественного отбора. Главная причина доминантности науки в нашей культуре состоит в том, что, осознавая наличие источников ошибок в самих себе и в группах, к которым они принадлежат, ученые разработали системы, связывающие их собственные руки. Сформировалась традиция критики, заставляющая всегда, при обнаружении изъянов, производить ревизию и усовершенствование.

Нам хотелось бы, чтобы наука была объективной, говорит Деннет, но нам не следует стремиться к нейтральности в науке. Есть плохая наука и есть ложные ее пути, но в целом все это не делает дороги науки чем-то менее объективным. Ирония состоит в том, что плоды научной рефлексии, свидетельствующие о темных пятнах несовершенства, временами используются теми, кто подозрителен к науке как их основе с целью отрицания ее привилегированного статуса в департаменте поиска истины — как будто институты и практики, в которых им видятся конкуренты науки, являются в этом отношении лучше.

Очень часто в самокритике науки видят свидетельство фундаментального изъяна самой науки и тщетности всех заявок на знание по причине релятивности последнего. К сожалению, в последнее время наука на этой основе все больше становится объектом политических злоупотреблений. Особенно опасно, что такого рода предубеждения против науки по-

360

лучили распространение в странах третьего мира. И здесь не нужно закрывать глаза на то, что свой вклад в мифологизацию науки внесли многие западные интеллектуальные течения, посеявшие здесь «вирус постмодернизма». Под действием этого вируса философия редуцируется к языковой игре или просто коммуникативному «разговору». К великому сожалению, среди людей, вносящих лепту в фривольное представление о философии, находится Рорти.

Критический вопрос о том, быть или не быть философии рационалистической деятельностью, напрямую связан с глобальным вопросом о контекстуальности или универсальности «западной»

рациональности. На этот счет у Деннета и Рорти разные мнения. Рорти — защитник идеи *солидарности*, которая включает в себя коммунитаризм, этноцентризм и культурный плюрализм. Это означает передачу понимания истины и ее критериев на суд любого сообщества, согласно его обычаям, ментальности и культуре. Главное здесь — служение солидарности, а не объективности. То, что истинно для западного мира, не является истиной для восточного, там все другое — другие понятия и другие критерии. Деннет — *защитник универсализма* «западной» науки и сформировавшихся в ее рамках принципов рациональности. Мой взгляд на науку, говорит он, соответствует взглядам Просвещения. И нет никаких оснований подвергать сомнению нормы западного рационализма и стандарты научной онтологии и эпистемологии, например, журнала «Тайм». Рорти же в этом видится местечковая платформа и «европоцентризм»<sup>1</sup>.

\* \* \*

Завершая сравнительный анализ позиций Деннета и Рорти, мы выражаем надежду, что экспликация их полемики прояснила какие-то важные моменты ситуации в современной

<sup>2</sup> [Dennett D. Back from Drawing Board // Dennett and His Critics. P. 205.](#)

361

англоязычной философии, не говоря уже о том, что в споре рельефнее высвечивается теоретическая суть контрпозиций. Poleмика поучительна и в идеологическом отношении: по ней в какой-то мере можно судить о напряжении, образовавшемся между расходящимися полюсами философии в США. Спор Деннета и Рорти перерос частности проблемы сознания и реализма и превратился в идеологический спор об ориентирах философии. Должна ли она стремиться к теоретизму и объективизму или согласиться с «литературным» статусом? В конечном счете, речь идет о стратегии философии: быть или не быть рациональной теоретической деятельности.

И еще. В отечественной литературе порою можно встретить констатации чуть ли не о доминантности постмодернизма в современной философской ментальности. Нам кажется, что это преувеличение. В заключении очерка о прагматизме Ричарда Рорти мы писали: «С нашей точки зрения, творчество Рорти следует оценивать по русской пословице: "На то и щука в море, чтобы карась не дремал". Его заслуга видится не в том, что он вызвал у кого-то растерянность и неуверенность в задачах философии, а в том, что предложенный им провокативный образ философии побудил философов-рационалистов продумывать заново посыпки, искать контраргументы, обновлять свой арсенал»<sup>1</sup>. Для историка философии, интересующегося динамикой ментальности современной философии, растущее сопротивление релятивизму и постмодернизму может служить хорошей иллюстрацией переменчивости баррикад, разделяющих разные позиции. Многие философы, разошедшиеся в разные стороны в толковании сознания, в ходе полемики вполне могут собраться по одну сторону баррикад, когда речь заходит о выборе позиции между рационализмом, объективностью и истиной, с одной стороны, и релятивизмом, ирреализмом и антитеоретизмом, с другой. Несмотря на существенные теоретические разногласия, Деннет, Сёрль, Нагель,

<sup>1</sup> Юлина Н.С. Постмодернистский прагматизм Ричарда Рорти. Долгопрудный, 1998. С. 98.

362

Чэлмерс в выборе философской идеологии объективно оказались по одну сторону баррикад, а Рорти — по другую. Можно даже сказать, что формируется своего рода фронт для отпора постмодернистскому вызову Западной Рационалистической Традиции, на которой возводилась Западная цивилизация и без которой невозможно Образование как Просвещение. К настоящему времени накопилось много свидетельств о набирающем силу отпоре постмодернизму (и социалистическому) со стороны защитников рационализма. С критикой постмодернизма выступили Томас Нагель, Эдвард Уилсон, Джин Брикмонт, Берн Буллох, Т. Шлик, Х. Сигель и др. Некоторые аналитические философы, до того достаточно индифферентно реагировавшие на его эксцессы, выступают с открытой полемикой, например, Дж. Сёрль<sup>1</sup>.

<sup>1</sup> См., например, полемику в 80-е годы между Сёрлем и Деррида: Searle John. Reiterating the Differences: A reply to Derrida // «Glypb», 1977; Derrida Jacques. Limited Inc. Northwestern Un. Press, Evanston, II. First published 1988, Third paperback printing, 1993. Четвертый раздел «Limited

Inc a b c... P. 29—110 содержит ответ на критику Сёрля; Сёрль Дж. Рациональность и реализм: что поставлено на карту // Путь. С. 212. Главный пафос книги Томаса Нагеля «The Last Word» (1997) — критика релятивизма и замены понятия «истина» понятием «солидарность».

## **У. К. Поппер и Д. Деннет: архитектура сознания согласно «Открытой» и «Закрытой» Вселенной**

Отрицая эссенциализм, самость все же можно описывать как «квази-сущность», как нечто, что представляется существенным для единства и непрерывности ответственной личности.

*Карл Поппер*

В 1977 году вышла в свет книга Карла Поппера и Джона Экклза «Самость и ее мозг. Аргумент в пользу интеракционизма»<sup>1</sup>, на которую последовало множество рецензий. Одной из самых резких была рецензия Дэниела Деннета<sup>2</sup>. Рецензент упрекал авторов во многих грехах, в том числе за то, что, сделав мишенью для атак старый редуционистский физикализм, они оставили без внимания принципиально новые нередуционистские (прежде всего функционалистские) материалистические теории. (Диалог Поппера с Экклзом он сравнил с «разговором олимпийских богов», не ведающих о том, что творится на земле.) В глазах молодого Деннета (тогда ему было 37 лет), полного энтузиазма относительно решения (или снятия) извечных головоломок сознания с помощью новых когнитивных средств — аналогий с

<sup>1</sup> Popper K., Eccles J. *The Self and Its Brain. An Argument for Interactionism.* Berlin, N.Y., L., 1977.

<sup>2</sup> *The Journal of Philosophy.* V. LXXVI. № 2. Febr. 1979. P. 91—97.

364

искусственным интеллектом и других, реабилитация авторами идей картезианства и интеракционизма выглядела, по меньшей мере, архаизмом. Ни Поппер, ни Экклз, насколько нам известно, никак не отреагировали на эту рецензию. Тем не менее имеет смысл сопоставить взгляды Поппера и Деннета на сознание. Сравнение позиций этих двух философов может пролить дополнительный свет на сложившуюся в современной философии сознания оппозицию монизма и дуализма, а также разницу стратегий, ориентирующихся, если прибегнуть к терминологии Поппера, на образы «Открытой» и «Закрытой» Вселенной.

Философию к Деннета и Поппера в принципе можно отнести к одному типу. Оба мыслителя являются натуралистами, эволюционистами, поклонниками науки, рационалистами, верящими в теоретизм философского знания. Оба работают в рамках «дарвиновской» процессуальной парадигмы, что не могло не повлиять на сходство их исследовательских стратегий и идей. Очень много «фамильных черт» можно найти в их эпистемологических установках. Читатель мог заметить, что в своем главном эволюционистском аргументе, в понимании роли «отбора» и «адапционизма» Деннет повторяет многое из сказанного Поппером о биологическом фундаменте знания, о приложимости к развитию знания принципа фальсификационизма и др. Напомним, Поппер утверждал, что в основе человеческого знания лежит «знание» биологических организмов и по отношению к любому знанию действует «естественный отбор», направленный на выживание более истинного в конкуренции с менее истинным (по принципу «try and eliminate errors»). Каждая «эмерджентия», будь то биологический организм или научная гипотеза, — это своего рода «теория» или «структура—ожидание» с заявкой на жизненность. Через пробные мутации и «элиминацию ошибок» происходит отбор наиболее приспособленных структур. В процессе адаптации к окружающей среде биологические виды приобретают истинное знание о мире (в противном случае они были бы элиминированы естественным отбором). Хотя их истинность, конеч-

365

но, не следует понимать в смысле достоверности (последняя приписывается только математическим истинам)<sup>1</sup>.

То есть и Деннет и Поппер подводят под знание биологические основы, верят в действие естественного отбора истинностных вариантов знания, рассматривают язык в рамках общего эволюционного развития. В целом можно сказать, что оба работают в рамках социобиологической

парадигмы. Эти сходные установки не могли не сказаться на сходстве подходов к сознанию. В российской литературе Поппер известен главным образом по его работам в области философии науки. В меньшей мере известна его позиция по проблеме сознания. Коротко напомним ее. Но сначала немного истории. В период разработки им концепции «Объективного знания» и концепции «Трех миров» (60-е годы) его внимание привлекли развернувшиеся в англоязычной философии дискуссии об отношении тела и сознания, самости, свободы воли. В доминировавшей в них тенденции к редукционалистскому физикализму и отказу от идеи реальности ментального он увидел влияние идей его коллег Р. Карнапа, У. Куайна, Г. Фейгла, с которыми у него был давний спор. В физикалистской тенденции Поппер увидел угрозу его концепции объективного знания и «Открытой Вселенной». Нужно сказать, что космологическая метафизика, которую он фактически выстраивал в результате размышлений о физике, математике, эволюционной теории, существенно отличалась от жесткой, нивелирующей качество, метафизики физикалистов. Приняв на вооружение идеи дарвиновского эволюционизма, эмерджентизма и процессуальности, он стремился нарисовать картину бытия без потери его качественного многообразия. Такая картина предполагает определенную целостность и обязательно должна включать феномен сознания. Поскольку этот феномен не укладывался в физическое объяснение, он не остановился перед тем, чтобы

<sup>1</sup> Подробно об этом см.: Юлина Н.С. Карл Поппер // Философы XX века. М., 1999. С. 167—187.

366

подвергнуть сомнению стандарты и законы существующей науки. Одним словом, толкование Поппером сознания и самости как реальных ментальных феноменов являлось важной составляющей концепции «Открытой Вселенной» и брошенного им вызова концепциям «Закрытой Вселенной».

В 1969 году он совершил поездку в США для чтения лекций в университете Эмори; в книжном варианте они названы «Знание и проблема тело—сознание. В защиту интеракционизма»<sup>1</sup>. В них он открыто объявил себя сторонником картезианства. Вопросы, которые ему задавали на лекциях, показали, что обсуждение психофизической проблемы требует осведомленности в нейрофизиологии. Не будучи специалистом в этой области, он вступил в творческий контакт с близким ему по взглядам Джоном Экклзом, нейробиологом с мировым именем. В 1974 году в Италии на вилле Сербеллони они провели серию бесед о сознании, мозге и множестве других проблем. Результатом их бесед явилась публикация совместной книги «Самость и ее мозг», название которой указывает на главную идею — именно самость является «владелицей» мозга, а не наоборот. В этой книге авторы, один с помощью философских аргументов, другой — научных, сформулировали основные тезисы, направленные против физикализма. Физикализм, толкующий единство знания на основе детерминистского панобъективизма, говорит Поппер, не в состоянии справиться с самым главным — с качественной спецификой содержания сознания. Приписывая статус реальности только физическому миру и заменяя сознание знаково-языковой коммуникацией и внешним поведением, физикализм упрощает картину бытия и недооценивает потенции природы к порождению принципиально нового. И сознание, и человеческая свобода, конечно, являются частью природы, вместе с тем, будучи качественно новыми образованиями, они трансцендируют ее. Вывод Поппера таков: физикализм исходит из

<sup>1</sup> Popper K. Knowledge and Body-Mind Problem. In Defense of Interactionism. L., N.Y., 1994.

367

концепции каузально «Закрытой Вселенной», между тем как «Вселенная Открыта»: она и детерминистична и индетерминистична, в ней возможны случайности и появление новых свойств или эмерджентий, каковым и является сознание.

Как дарвинистов Поппера и Деннета роднит *когнитивизм и процессуальность* в подходе к сознанию. (Термином «когнитивизм» в современной философии обозначают позиции, делающие акцент на содержательной и логической сторонах сознания.) Оба отрицают возможность объяснения сознания на основе психофизической корреляции. Согласно Попперу, хотя нейрофизиологические процессы мозга происходят одновременно с сознательными процессами,

первые нельзя считать презентациями содержательных объектов мышления. Идея бесконечности числового ряда, например, не может быть объяснена физическими событиями мозга, происходящими при размышлении о потенциальной бесконечности. Она принадлежит сфере идеального, в его терминологии — Миру 3. Главное в сознании — оперирование с помощью языка и семиотических средств содержательными единицами: проблемами, теориями, логическими следствиями из посылок и т.п. Однажды возникнув, эти единицы способны к *автономному* развитию (в изобретении числового ряда содержались потенции развития, приведшие к современной математике). В отличие от физического мира (Мира 1), у Мира 3 нет пространственно-временных характеристик.

Примерно в том же духе вопрос о соотношении нейрофизиологии и содержания сознания решает Деннет. Он не исключает, но считает маловероятным объяснение содержания сознания нейрофизиологическим устройством мозга. Секрет сознания состоит не в наличии в физическом мозгу связки нейронов, а в том, что сознание представляет собой чистую функцию, логически упорядоченные информационные процессы, в которых *мысли сами себя мыслят*: из одной идеи рождается другая, из нее третья и т.д. Мысли всегда представлены в виде языково оформленных единиц. Как и Мир 3 Поппера, содержательная составляющая когнитивного процес-

368

са, в терминологии Деннета «Информационная диаспора», не имеет пространственно-временной локализации, она — *виртуальная машина* без свойств.

Но дальше начинается существенное расхождение. Поппер без колебания назвал Мир 3 «реальностью», существующей наряду с физическим миром. Деннет категорически против умножения «реальностей». Виртуальная машина сознания — это только функциональные отношения, а не особое свойство, имеющее статус реальности: кроме физической, никакой другой реальности не существует. Тем не менее, если не обращать внимание на разницу философского различия понятий «виртуальный» и «реальный», когнитивизму этих двух мыслителей, с нашей точки зрения, присущи признаки платонизма: в любом толковании содержание сознания выносятся на иной, по сравнению с телом, уровень.

Все же главное расхождение Деннета и Поппера начинается в другом — там, где заходит речь об «авторе» мыслительных процессов: выступает ли их автором человеческая личность (самость) или они могут быть безличными. По Попперу, процесс мышления не может происходить без «автора». Именно самость (или сознание живой человеческой личности) является «автором», направляющим, контролирующим и корректирующим все сознательные процессы. Самость, говорил он, столь же реальна, как столы и стулья, но в отличие от последних обладает свойством самоактивности. Благодаря ее самоактивности творятся смыслы, а вместе с ними артефакты и культура. Ее роль он сравнивал с ролью программиста, если мозг сравнивать с компьютером. «Самости являются единственными активными агентами во Вселенной: единственными агентами, по отношению к которым термин «активность» может быть применен в подлинном смысле этого слова»<sup>1</sup>.

У Деннета, как мы видели, совершенно иной взгляд на авторство. Для того чтобы мысли себя мыслили, а одни тексты порождали другие тексты, «автор» не обязательно должен

<sup>1</sup> Popper K., Eccles J. The Self and its Brain. P. 538.

369

быть человеческим субъектом. Там, где мысли сами себя мыслят, не имеет значения, в ком или в чем реализуется мыслительный процесс. Вполне возможна *множественная реализация*: объектом реализации может быть органический мозг, физический компьютер или социальный институт. Производимая компьютером новая информация тоже может интерпретироваться как самотворчество; она не сводится к физическому веществу, устройству или исходной программе компьютера. Представление об уникальной природе ментальной самости, командующей сознательными процессами, пришло в противоречие с фактом когнитивных возможностей компьютера. Поэтому утверждать, что только человеческий субъект способен к творчеству, — ошибочно. В принципе, робот тоже способен к саморефлексии. То, что мы представляем в виде



наших интимных самостей, в действительности сплетено из имеющихся в социуме дискурсов нашими разговорами о самих себе и разговорами о нас других людей, то есть является артефактами культуры, а не проявлением метафизической самости.

Разное понимание авторства сознания, естественно, сказалось и на разном понимании его «архитектуры». Согласно Попперу, в «архитектуре» сознания присутствуют три уровня бытия или три «реалии»: нейрофизиология (Мир 1), субъективные ментальные состояния (Мир 2) и объективные содержательные продукты мышления (проблемы, теории, идеи и другие «идеальные обитатели» Мира 3). Напомним, что при объяснении объективности знания Поппер взял на вооружение антипсихологический лозунг «Познание без познающего субъекта». Тем не менее, столкнувшись с необходимостью объяснения феномена сознания, он ввел в свою схему Мир 2, роль которого состоит в осуществлении интеракции между физическим миром (Миром 1) и миром идеальных форм (Миром 3). Без введения такого медиума — субъективных, психологических, ментальных состояний — он не видел никакой другой возможности объяснить взаимодействие нейрофизиологии с идеальным содержанием Мира 3. (С нашей точки зрения, Мир 2 у Поппе-

370

ра прописан слабее Мира 3. Остается неясным, откуда у Мира 2 появляется способность к усвоению языка и интеракции «миров» и отчего он обладает чудесным свойством наделять самость свободой и самоактивностью.)

Слоган «познание без познающего субъекта», как нам кажется, с большим правом применим именно к позиции Деннета. Деннет прекрасно понимает уязвимость любого психологизма (и играет на этом). Он принимает Мир 3 (как когнитивный, информационный срез сознания, но не как «реальность»), при этом элиминирует из «архитектуры» Мир 2 — мир субъективного и феноменального. Допущение его реальности, убежден он, блокирует объяснение сознания.

Попперовская идея «трех реальностей» с точки зрения Деннета — неоправданная спекуляция. Реальность только одна — физическая; исходя из нее, мы должны объяснять все явления. Ментального посредника между нейрофизиологией и содержанием мышления не существует. Представление о его существовании — картезианская иллюзия, возникшая из нерелексивной интуиции. Сознание есть функция без субстрата, отношение без свойств, виртуальная машина, запущенная процессом коэволюции биологического и социального и имплементированная в биологии мозга.

В свете широкого применения Деннетом компьютерных аналогий интересно посмотреть, как Поппер отнесся к феномену искусственного интеллекта (AI). (У него не было специальной работы на этот счет, однако отдельные высказывания говорят о наличии у него ясного мнения.) Если использовать нынешнюю классификацию позиций в отношении искусственного интеллекта, Поппера, наверное, можно было бы отнести к сторонникам «слабого AI». Как мы уже говорили в предыдущих разделах, сторонники «слабого AI» считают возможным симуляцию работы мозга (и сознания) на компьютере, но отрицают дублирование; субъективность и феноменальность опыта не поддаются этому.

Поппер назвал компьютер «типичным объектом Мира 3» и в смысле его проекта, и в смысле производимой им продук-

371

ции — информации. Он говорил следующее: «Хорошо натренированный мозг, как и надежно работающий компьютер, работают в соответствии с принципами логики, присущими Миру 3, *так же*, как и в соответствии с законами физики и электрохимии»<sup>1</sup>. Стандарты логики являются бестелесными стандартами. На таких же стандартах строятся и программы компьютеров, выступая стандартами достоверности и рациональности. Появление новых технических средств делает возможным решать проблемы более эффективно. Нет никаких оснований считать компьютеры глупее нас. «Эйнштейн как-то сказал: "Мой карандаш умнее меня". Под этим он, конечно, подразумевал следующее: при помощи карандаша можно получать результаты, которые без него нельзя представить. Компьютер не что иное, как "прославленный карандаш": вооруженные им, мы можем стать более чем в два раза умнее, чем без него. Вооруженные компьютером (типичным

объектом Мира 3), мы, возможно, станем в сотни раз умнее, чем без него; с более совершенными компьютерами нет нужды ставить этому границы»<sup>2</sup>. Тем не менее заявления энтузиастов, что наши мозги, а может быть, и наши сознания, являются компьютерами, Поппер отнес к фантастике. При всем глубочайшем уважении к Тьюрингу, говорил он, я не могу согласиться с ним, что компьютеры могут *мыслить*. «"Разум", который усматривают в компьютере программисты или специалисты по искусственному интеллекту, вложен в него нами, и то обстоятельство, что компьютер может делать *больше* того, что делаем мы, обязано факту, что именно мы вкладываем в компьютер мощные операционные силы; на деле — принципы автономного Мира 3»<sup>3</sup>. Наши самости создали идею компьютера и придали значения символам программ, а не наоборот. Как мы уже сказали, и Поппер и Деннет в общем и целом работают в рамках социобиологической парадигмы. Внутри ее

<sup>1</sup> Popper K., Eccles J. *The Self and its Brain*. P. 77.

<sup>2</sup> Ibid. P. 208.

<sup>3</sup> Ibid. P. 208, сноска 3.

372

возможны различные стратегии. При всем том, что у Поппера Мир 3 состоит из языковых и семиотических единиц (т.е. социологичен), в объяснении сознания он в большей мере склонен к биологизму (в этом отношении он близок Дж. Сёрлю, А. Кларку, П. Черчленду). Он не исключал, что люди когда-либо смогут научить шимпанзе хотя бы рудиментарно мыслить и что углубление знания принципов естественного отбора поможет с применением искусственного отбора создать какие-то новые биологические виды, способные конкурировать с людьми. Во всяком случае, заявлять о невозможности всего этого было бы неосмотрительным. Однако Поппер был твердо убежден, что «мы не в состоянии построить электронные компьютеры, наделенные сознательным субъективным опытом»<sup>1</sup>. У Деннета противоположное мнение. Если главным признаком сознания считать функционирование когнитивных информационных процессов (то есть признак, присущий Миру 3 Поппера), тогда привязка их к биологии становится не обязательной и пропасть между сознанием и компьютером исчезает. Начиненный гигантской информацией робот в принципе мог бы дублировать человеческое сознание: и не только в смысле логических операций, но и в смысле самоинтерпретации собственных мыслей. В принципе, говорит он, «могла бы существовать сознательная самость, чьим телом был бы робот, а мозгом компьютер»<sup>2</sup>.

Как всегда, спор философов о «что» упирается в вопрос «как»: как обосновать достоверность принятых стратегий? По Деннету, нет никаких оснований подвергать сомнению принятую в науке каузальную (и интерсубъективную) модель объективности. Нет оснований думать, что идея эмерджентизма несовместима с детерминизмом, как нет смысла гадать о будущем открытии новых физических законов, охватывающих и атомы, и радость от хорошей погоды, или уповать на появление принципиально нового концептуального аппарата (что

<sup>1</sup> Ibidem.

<sup>2</sup> Dennett D.C. *Consciousness Explained*. L., Boston. 1991. P. 431.

373

свойственно Дж. Сёрлю, Т. Нагелю, К. МакГинну и другим пишущим о сознании современным философам). В мире действует принцип каузальности и, используя терминологию Поппера, «Вселенная Закрыта». Нужно только с большим философским воображением использовать зарекомендовавшие себя в естественных науках модели объяснения. Уже сегодня они дают возможность нарисовать *нейтральную картину мира*, в которой нет необходимости делить ее на «реальности».

Поппер, конечно, понимал, что его тезис о реальности сознания и самоактивной самости вступает в противоречие с физическими законами термодинамики. Преодоление этого противоречия он связывал с революцией в физике и принятием гипотезы «Открытой Вселенной». Если физика не в состоянии найти в своих законах место для сознания, то, как говорится, тем хуже для нее. Возможно, что в будущей физике такое место найдется. Здесь мы хотели бы подчеркнуть, что в

отличие от Нагеля, Сёрля и других философов, абстрактно уповающих на будущую концептуальную революцию, Поппер более конкретен. Он связал революцию в науке с принятием гипотезы «Мира предрасположенностей» или диспозиций, не поддающихся детерминистскому объяснению и вероятностному исчислению<sup>1</sup>. Эта гипотеза, считал он, помимо того, что устраняет многие противоречия теории вероятности, дает основу для объяснения сознания и свободы человека. Примечательно при этом, что Поппер не выводит «предрасположенности» за рамки физики. Он относит их к реальным «ненаблюдаемым диспозиционным свойствам физического мира» или «полям физических сил». Иначе говоря, вопреки прокламируемому им дуализму и поношениям физикализма, конечное (идеальное) объяснение сознания он мыслил все же физикалистски и в этом отношении мало чем отличался от критикуемого им Куайна.<sup>2</sup> Одним словом,

<sup>1</sup> Popper K. A World of Propensities. Bristol, 1990.

<sup>2</sup> Поппер. К. Логика и рост научного знания. М., 1983. С. 414—438. Подробнее об этом см.: Юлина Н.С. К. Поппер: мир предрасположенностей и активность самости // Философские исследования. 1997. № 4.

374

понимание им как дуализма, так и физикалистского монизма еще очень неопределенно. Сказанное относится не только к Попперу. Нужно заметить, что, несмотря на широкое употребление в современной философии термина «физикализм», смысл его еще далек от ясности. Чаще всего физике приписывается онтологический авторитет относительно того, что есть в мире, и эпистемический авторитет в отношении стандартов получения адекватного знания о мире. В таком толковании остается много неясного. Понимается ли под физикализмом «каузальная закрытость» физического мира, то есть все, о чем может быть сказано как о существующем, может быть только физическим и подчиняться законам физики? Или физическая система, говоря словами Поппера, является «открытой»; одновременно и детерминистской и индетерминистской, законосообразной и случайной, способной порождать непредсказуемые, не укладывающиеся в традиционные каузальные схемы феномены? Если имеется в виду «закрытость», можно ли считать язык физики адекватным для описания всего сущего, в том числе фактов, изучаемых в биологических, психологических и социальных дисциплинах, использующих другие языки, не редуцируемые к языку физики? А если это не так, почему их языки нельзя признать адекватными для описания того, что есть? в том числе признать язык психологии адекватным для описания ментальных явлений? Возникают и другие вопросы. Исчерпывается ли содержание физикализма физикой наших дней или физикализм имеет в виду какую-то будущую идеальную физику? Если имеется в виду первое, то это неверно. Вряд ли можно отрицать, что нынешняя физика неполна, развивается и со временем ее содержание может существенно измениться, как это произошло в движении науки от Ньютона к Эйнштейну. Если же имеется в виду второе, то никто не знает, что станет содержанием будущей идеальной физики, и не исключено, что там появятся какие-то принципиально новые законы. Во всяком случае, нельзя утверждать с уверенностью, что в ней не образуется законное пространство для «духа». На эти вопро-

375

сы физикалисты обычно отвечают так. Со временем, конечно, физическое знание изменится. Возможно, что для объяснения некоторых феноменов в физику потребуется ввести нередуцируемые «духи», но тогда «духи» будут физическими по определению. Однако ни сейчас, ни в будущем не будет иного, кроме физики, стандарта, в соответствии с которым можно было бы наиболее адекватно отвечать на вопрос о том, что есть в мире.

Со времени выхода в свет книги Поппера и Экклза «Самость и ее мозг» и последующей на нее рецензии Деннета прошла четверть века. Обозревая проходившие за это время дискуссии о сознании, можно сказать, что трудности в объяснении человеческого сознания и субъективности (при помощи компьютерных метафор или без них) заставили многих философов дрейфовать к различным версиям дуализма. Мы уже указали на некоторые переклички между позициями Поппера и Джона Сёрля: оба связывают сознание только с одним органическим носителем —

мозгом; инкарнация сознания в разных носителях невозможна. С другой стороны, сторонники физикалистского монизма, опираясь на растущий массив данных когнитивных наук о ложности интроспективных отчетов субъектов о своих внутренних состояниях, усиливают атаки на феноменализм: они настаивают на социальности сознания, его творимости языком и культурой и отсюда о принципиальной возможности его компьютерного дублирования. Одним словом, вопрос о сознании остается открытым, а спор дуалистов и монистов продолжается.

И последнее историко-философское наблюдение. Сходство позиций этих двух мыслителей мы видим еще и в том, что и тот и другой начали со скромных «инженерных» заявок, а пришли к Grand Syntheses. Подобно тому как в своем синтезе Поппер остановился перед вопросом, относящемуся к Misterium Tremendum, — «Почему мир таков, какой он есть?», Деннет тоже проявляет нерешительность перед «конечным вопросом» (Ultimate Question) — «Почему Матери-Природе свойственны исходная интенциональность и талант дизайнера?»

## Заключение

### Заметки историка философии

В начале нашего рассказа о Деннете мы обещали читателю представить его прежде всего как современного мыслителя, выступившего с грандиозным проектом по устранению с пути материалистического монизма самого главного препятствия — тайны сознания, с тем чтобы навсегда покончить с дуализмом. Снять головоломки, возникшие вокруг этого загадочного феномена таким образом, чтобы это соответствовало физикалистским интенциям науки и в то же время сохраняло самые лелеемые человеком ценности — самость, свободу, личностное, моральность, ответственность. По окончании нашего рассказа уместно спросить: удалось ли Деннету представить сознание «объясненным»? в чем состоит, если таковой есть, его вклад в философию сознания? что представляет собой (как оценить) сам феномен философии Деннета?

На первый вопрос у нас нет однозначного ответа. Нам кажется, что тайна сознания слишком велика, чтобы с ней можно было покончить с помощью софистичной, вооруженной новейшими познавательными инструментами теории. Как и при исследовании загадок материи, за частичными успехами обнажается новый, еще более запутанный клубок загадок. Каждое продвижение в понимании феномена осознающего себя интеллекта натывается на слишком грандиозную тайну: как вещество мозга порождает мысли, материальный (невербальный) уровень переходит на вербальный, а биологическое стыкуется с социальным? Разумеется, для ее раскрытия тре-

377

буются неординарные философские идеи, методы, понятия, инструменты. Но и всего этого тоже недостаточно. Требуются еще очень большие усилия ученых по исследованию физической и биологической материи, новый уровень междисциплинарных исследований, охватывающих такие разные области, как теория эволюции, теория искусственного интеллекта, биоинформатика, нейробиология, психология, когнитивные науки, психолингвистика, психофармакология, психосемантика и другие.

Субъективно нам не очень нравится позиция убежденных агностиков и пессимистов, заявляющих, что ответ на вопрос «что такое сознание?» — вне возможностей человека. Пессимизм — не самый лучший настрой при исследовании чего угодно. На его фоне оптимизм Деннета выглядит привлекательнее; плодотворные и новаторские идеи в истории знания чаще всего возникали у оптимистов. Тем не менее аргументы Деннета не смогли подавить сомнения. Нам кажется, что, когда умеренные пессимисты связывают ответ на этот ключевой вопрос с грядущей концептуальной революцией и радикальной ломкой наших представлений о материи, для этого есть основания. Хотя, если посмотреть с другой стороны, Деннет прав, говоря, что пассивное упование на прогресс в неопределенно далеком будущем — не самая лучшая позиция. Любые загадки следует разгадывать имеющимися в наличии средствами.

Несмотря на наши сомнения относительно состоявшегося перевода сознания из разряда «загадок» в разряд «отгадок», в целом мы оцениваем проект Деннета как *успешный, как новизну и вклад* в философию сознания. Вынося такой вердикт, нужно пояснить критерии, которые имеются в виду. Что вообще значит успешность, новизна и вклад в философию? Этот вопрос может иметь два смысла: первый — решена ли какая-либо проблема (раскрыта ли окончательно тайна сознания); второй — есть ли продвижение к раскрытию тайны? Историк мысли, которого интересует развитие философии в контексте новаций культуры, скорее всего отрицательно ответит на первый вопрос и положительно на второй.

378

Томас Кун говорил, что научные парадигмы различаются между собой не столько ответами, сколько вопросами. Именно в характере последних, прежде всего, сказываются сдвиги когнитивной культуры. Этот тезис можно отнести и к философии, хотя здесь имеются свои особенности, связанные с тем, что новые вопросы в каждую последующую эпоху соотносятся с



«вечными проблемами». С точки зрения историка мысли, прогрессивное движение в философии (да и в любом виде знания) может проявляться в различных вещах, не обязательно в «окончательных решениях». Прогресс имеет место тогда, когда в исследуемом объекте проводятся более *тонкие дистинкции*, в результате чего в нем обнаруживаются новые, неожиданные аспекты. Как правило, это приводит к постановке вопросов, которые раньше никому не приходили в голову. В итоге появляется множество *подпроблем* и новых понятий, в свете которых переформулируются старые проблемы. Иначе говоря, происходит то, что Пол Фейерабенд назвал «пролиферацией» идей. Все это находит свое выражение в *новом языке* (или специфическом для данной области жаргоне), появление которого, безусловно, следует отнести к очень важной новации. Если исходить из общего понимания прогресса культуры как расширения смыслового поля языка и умножения семиотических средств, тогда все это означает появление в философии новых понятийных единиц, иначе говоря, *расширение смыслового поля философии*. Другой важный показатель — насколько в содержании и в форме обсуждаемых вопросов имеют место *когнитивные сдвиги под влиянием* введения *нового инструментария*, позволяющего проводить более тонкую, разнообразную и грамотную работу. Например, введение в арсенал философии техники логического и концептуального анализа было шагом вперед в ее когнитивной культуре. То же самое можно сказать о введении компьютерных метафор и аналогий: как бы их ни оценивать, они стали новыми мощными инструментами, помогающими понять многое и в «архитектуре» сознания, и в работе его «машины». Именно в свете этих критериев, как нам кажется, следует оценивать проект Деннета по объяснению сознания.

379

Вклад Деннета в исследование сознания условно можно подразделить на две части — критическую и конструктивную. Критическая часть состоит в терапии принятого в философии сознания языка и тематики, в прояснении, уточнении и обновлении проблематики, относящейся к феномену сознания, которая со времен Декарта оставалась достаточно смутной. Более определенным стало поле исследования, систематизированы головоломки, над разгадыванием которых бились многие философы прошлого, произведен их смысловой анализ и показано, какие из них достойны разгадывания, а какие являются мнимыми. Хотя многие оспаривают его диагноз и предложенную терапию, сам проведенный Деннетом анализ — это шаг вперед. Главное в нем, конечно, состоит в показе противоречий менталистских и феноменалистических концепций и в выявлении изъянов представлений о сознании как состоящем из дискретных, непосредственно наблюдаемых в опыте качеств. Деннет высказал ряд заслуживающих внимания доводов в пользу предположения, что, поскольку объяснение сознания в рамках единства мира на основе предпосылки о реальности феноменальных черт опыта невозможно, нужно искать радикально новое объяснение опыту и субъективности. Как бы ни относиться к предложенному им методу гетерофеноменологии, в свете его виднее уязвимые места феноменологии. Его аргументы против дуализма и феноменализма дали богатую пищу дискуссиям и стимулировали новые поиски объяснения когнитивных и чувственных сторон опыта человека.

Конструктивная сторона философии Деннета тоже достаточно весома. Им выдвинут ряд оригинальных предположений относительно возможных решений проблемы сознания, если сравнивать их не только с предшествующими, но и с нынешними проектами. В обиход были введены несколько новых концептуальных схем, идей, понятий и метафор: Модель Множественных Набросков, таксономия интенциональных установок, виртуальная машина сознания, язык как кодировщик мозга, гетерофеноменология, самость как «центр

380

нарративной гравитации», алгоритмы «Машины Дарвина», «Машины Тьюринга», «Машины Сознания», различение *abstracta* и *illata*, идеи «реальных паттернов», «свободно плавающей рациональности» и многое другое.

Одна из важных новаций Деннета — Модель Множественных Набросков. Выдвинутая как альтернатива традиционному «гомункулярному» образу сознания, в котором главным детерминантом и «означивателем» выступает самоактивная самость, он рисует

децентрализованную картину сознания, основанную на «мультитрековой работе мозга». Основное место в ней занимает не отношение нейрофизиологии и духа, а информационное производство, деятельность «виртуальной машины» в кодировке содержательной, когнитивной составляющей сознания. Один из самых дискуссионных вопросов в современной философии сознания — каким образом преодолевается пропасть между нейрофизиологией и феноменологией? Согласно его версии функционализма, информационные процессы (деятельность виртуальной машины) выполняют опосредующую (нейтральную) роль. Другой спорный вопрос, «что явилось первичным, а что вторичным в появлении сознания — социальное или биологическое?», Деннет решает, можно сказать, диалектически. В отличие от биологицистов, объясняющих генезис сознания случайными мутациями, инновациями (эмерженциями) природы, он отводит решающую роль в возникновении интеллекта искусственно созданным знаковым системам. Мозг становится органическим компьютером, когда его биологические информационные системы кодировок под влиянием знаковых систем языка и культуры трансформируются в новое эволюционно-культурное качество — сознание. Процессы кодировки, присущие информационным системам природы, на каком-то этапе эволюции только у вида *homo sapiens* превратились в знаковые системы, применимые в разных контекстах, а не только в тех, которые запрограммированы биологической эволюцией. Опираясь на ряд научных исследований последнего времени, Деннет выдвинул оригинальную гипотезу о влиянии семиотических си-

381

стем на органику мозга, в результате чего в параллельной работе мозга появились процессы, имеющие сериальный характер. Эта гипотеза широко обсуждается в философской и научной литературе. С его точки зрения, сознание не есть случайная биологическая инновация (мутация), оно сотворено одновременно с эволюцией знаковых систем языка и культуры.

К оригинальным новациям Деннета следует отнести теорию интенциональности. На горизонте любого физикалистского мировоззрения постоянной тучей висит «тезис Brentano» — пропасть между рисуемой наукой картиной мира, в которой в общем и целом действует принцип научной рациональности, и миром повседневных представлений, в котором действуют целеполагание, мораль, ответственность, нормы и ценности. Приняв «вызов Brentano», он предложил свою версию соединения «миров», построенную на «внешнем» или «приписывающем» понятии интенциональности. Понятие «интенциональность» несет у него очень большую эпистемологическую и метафизически коммуникативную нагрузку. Одна из главных задач — узаконить принцип рациональности в поведенческой, интенциональной сфере. (То, что он называет «интенциональной сферой», можно было бы назвать, используя термин Канта, сферой практического разума.) В отличие от физической установки и установки дизайнера, интенциональная установка включает в себя *интерпретацию* поведения объекта в соответствии с допущением действия принципа рациональности, то есть представление объекта как имеющего определенные цели и достаточную (хотя бы минимальную) рациональность для достижения этих целей.

Интерпретивизм, реляционизм, нормативизм — это типичные черты конструктивистской эпистемологии. Самая интересная часть эпистемологии (и метафизики) Деннета состоит в натурализации интерпретивизма. Он отстаивает тезис, что интерпретивизм не обязательно вступает в противоречие с реализмом и объективизмом при условии принятия эволюционного принципа и идеи протоинтенциональности Матери-Природы. Первичный источник интенционального находится в дизайне

382

самой природы. Деннет, конечно, не приписывает эволюционному принципу статус логического или гносеологического аргумента. Он настаивает на другом: если не принять его в качестве главного ориентира, невозможно объяснить рациональное поведение и знание. При этом он одновременно и натурализирует, и социологизирует интенциональность.

Одна из его новаций — идея «свободно плавающей» (free-floating) рациональности (и интенциональности). Ее смысл состоит в том, что действие естественного отбора может

осуществляться в самых разных системах, не обязательно имеющих сложные информационно-процессуальные возможности, например, в растениях, животных, компьютерах. С помощью таксономии установок он поднимает пронизанный культурой и «интенциями третьего порядка» мир человека высоко над биологической природой и одновременно погружает этот мир в природу, приписывая человеку, наряду с растениями и животными, интенции низшего уровня — первого и второго порядка. В принятии иерархии установок он видит единственный путь к снятию антагонизма механистического и целополагающего объяснений человека и к узакониванию статуса моральности и ответственности.

Заслуживает всяческого внимания стратегия Деннета на соединение специфически философского и научно-натуралистического подходов в объяснении феномена сознания. Даже для западной философии очень большая редкость, когда профессиональный философ решается предложить свое видение дарвинизма и эволюционизма и оспаривает позиции специалистов-биологов. Вызвавшая волну дебатов его книга «Опасная идея Дарвина» стимулировала дискуссии вокруг эволюционизма: ученых она заставила вникать в философские аспекты эволюционизма, а философов — обращать внимание на имеющиеся научные тонкости в понимании эволюционного принципа. И, может быть самое главное, она заострила внимание на тезисе, что без принятия этого принципа кардинальные (эпистемологические, метафизические, этические) вопросы философии неразрешимы.

Деннет не ограничился концептуальным анализом разговора о сознании, характерным для философии лингвистического ана-

383

лиза. С помощью эволюционного аргумента он попытался избежать ловушки социолингвистической парадигмы («языковой тюрьмы» Витгенштейна) и заключенной в ней тенденции к релятивизму. Работать в социолингвистической парадигме, как мы уже отмечали при сравнении позиций Деннета и Рорти, сравнительно легче, нежели в социобиологической парадигме. Роль социума и языка в формировании интеллектуальности сравнительно хорошо изучена (чего нельзя сказать о биологических диспозициях интеллекта). Однако социологи обходят ключевой вопрос о биологических корнях социальности. Заслуга Деннета состоит в том, что главный предмет исследования — сознание — он расположил на стыках биологического и социального, нацелившись перекинуть через них мосты. Вольно или невольно он вынужден балансировать между разными подходами, прибегать к двойной перспективе, что не может не быть сопряжено с непоследовательностью. (Нам кажется, что зарубежные комментаторы, критикующие его за непоследовательность, недооценивают трудности выбранной им перспективы.) Деннет не побоялся ринуться в неизвестное и строить гипотезы о появлении «мемосферы» как новом этапе эволюции, о воздействии культуры (мем-вирусов) и обучения на микроструктуру мозга, о сотворении сознания на стыке биологических и социальных кодировок, о децентрализованной работе мозга (Модель Множественных Набросков), о диалектике довербального и вербального, о появлении сериальных шагов в параллельной работе мозга и др. Его интерпретация на основе эволюционизма акта творчества (и создания артефактов культуры) — это новое слово в современной философии. О феномене творческой активности человека написаны горы литературы, в значительной мере спекулятивной. Деннет же попытался строить свои гипотезы в соответствии с имеющимися на сегодня научными данными и не покидая почвы рационализма.

Представляя философию Деннета, мы обещали показать его как «инженерно мыслящего» аналитика, у которого «минимальная метафизика» перерастает в «максимальную метафизику». По нашему наблюдению, многие комментаторы, а среди них

384

немало «глухих» к метафизике философов и ученых, упускают из виду этот важный момент. Во многом это объясняется нежеланием Деннета заниматься «высокими» сюжетами и его скептицизмом в отношении обобщений, в которых, как правило, смазываются сложности реальности. Хотя у него часто мелькает фраза «это не соответствует моим метафизическим

убеждениям», соединять эти убеждения в цельную систему он не стремится: к такого рода деятельности он испытывает идиосинкразию. И все же такая метафизика есть, хотя историк, описывающему его творчество, приходится ее реконструировать из рассуждений на разные темы. Ведь сама специфика проблематики сознания такова, что не может не выводить на глобальные космологические вопросы о природе мира, в котором в единстве сосуществует и физическое и психическое. В одном из разделов книги, сопоставляя его позицию с позицией Поппера, мы говорили о том, что молодого Деннета раздражала манера Поппера выводить проблему сознания на «высокий» космологический уровень. С годами у него самого проявилась схожая тенденция. Эволюционизм и натурализм — фамильные признаки процессуальной философии — с необходимостью толкают к глобалистике и метафизике.

Как мы уже говорили, Деннет избегает идентифицировать свою концепцию одним термином: в подведении под «изм» ему видится аристотелевский импульс схватить сущность в понятии. Свою теорию он именуется «процессуально-информационной теорией AI», «физикализмом», «функционализмом», «материализмом», «семиотическим материализмом», считая все их неадекватными. Комментаторы тоже именуют ее по-разному: «логическим бихевиоризмом», «компьютерным материализмом», «эволюционным когнитивизмом», «информационной теорией» и другими терминами. Учитывая, что наша работа является историко-философской, а в этом жанре не обойтись без увязки с традициями, историческим контекстом и типологизацией, нам кажется уместной ассоциация его философии с рядом более известных в российской литературе «измов» — нейтральным монизмом, сциентизмом, физикализмом.

385

Концепцию Деннета, как нам кажется, без большой натяжки можно отнести к традиции «нейтрального монизма». Его версия нейтрализма имеет отличительные особенности: чтобы увидеть их, полезно напомнить историю предпринимавшихся в XX веке восстаний против дуализма с позиции нейтрального монизма.

Все имевшиеся в истории мысли монистические концепции объединяла оппозиция дуализму и тезис о непрерывности и единстве мира. Этот тезис обосновывался либо с идеалистических, либо материалистических позиций. В первые десятилетия XX века попытки преодолеть оппозицию идеализм—материализм с помощью нейтрального монизма были предприняты эмпириками. Опыт толковался таким образом, при котором он редуцировался к чему-то «третьему», не являющемуся ни физическим, ни психическим. Как правило, этим «третьим» выступали «элементы опыта»<sup>1</sup>. Можно вспомнить концепции российских эмпириков — эмпириомонизм А.А. Богданова, эмпириосимволизм П.С. Юшкевича, философскую концепцию Н.В. Валентинова и др.<sup>2</sup>. На Западе нейтралит-

<sup>1</sup> А.А. Богданов писал: «Бесконечный поток опыта... представляет в своем целом не только очень грандиозную, но и очень пеструю картину. Разлагая шаг за шагом это целое, анализ переходит от более крупных его частей ко все более и более мелким и достигает наконец некоторой границы, где разложение дальше не удается. Здесь лежат *элементы опыта*». (Богданов А.А. Эмпириомонизм. М., 2003. С. 6.)

<sup>2</sup> Можно напомнить, что в «Материализме и эмпириокритицизме» В. Ленин негативно отзывался о дискуссиях российских эмпириокритиков, расценив их как свару «желтых и зеленых чертей». На самом деле в российской философии того времени шла напряженная работа по детализации и выработке нового языка в описания физической и психической сторон опыта и по обозначению проблематики, которая охватывается этими понятиями. Особо хотелось бы отметить, что вся эта работа велась в благоприятной для творчества дискуссионной форме. Были высказаны интересные идеи и гипотезы, которые не получили дальнейшего развития и не стали мировым достоянием в силу социальных и политических причин, прервавших эту дискуссию.

386

ские варианты монизма развивали Уильям Джемс, Ральф Перри, Джон Дьюи и многие другие. Наибольшую известность получил «нейтральный монизм» Бертрана Рассела. Новшество, которое он (а затем неопозитивисты) внес в нейтрализм, состояло в переводе разговора о единстве мира в плоскость языка. Представления о ментальном и физическом были поставлены в логическую

зависимость от высказываний о «событиях опыта», являющихся базисными и «нейтральными» в отношении их описания в физических или психических терминах. Рассел полагал, что при условии прояснения смысла и логических структур наших высказываний можно показать изоморфизм между структурой предложений и структурой мира. При этом ни он, ни неопозитивисты еще не порвали окончательно с феноменализмом: чувственный опыт выступал базисом теоретических конструкций.

Другое дело чуждый феноменализму логический бихевиоризм. Преодоление дуализма и достижение нейтрализма в этом случае достигаются путем выведения из объяснительной схемы ментального и феноменального и замены его бихевиоральным. Представление о целостности мира поставлено в зависимость от нашего поведения в практическом мире здравого смысла, в котором язык употребляется для коммуникативных и прагматических целей. Джилберт Райл, известный своей атакой «Дух в машине», считал, что мир един, однако судить об этом мы можем не на основе его физической или ментальной природы, а на основе нашего разговора о нем и нашей жизнедеятельности в нем. Ничего другого о его «сущности» сказать нельзя. По Райлу, бессмысленно называть себя материалистом (или идеалистом), когда сама оппозиция физического и ментального является «категориальной ошибкой».

В каком-то отношении Деннет продолжил линию логического бихевиоризма, но, в отличие от Райла, не боится причислять себя к лагерю материалистов, физикалистов, натуралистов. Суждение о единстве мира на лингвистической основе для него не есть подлинное объяснение: из него выпадает очень важный момент, что сам язык является результатом

387

развития глубинных материальных эволюционных процессов. Без обращения к ним онтологическая оппозиция ментального и физического не снимается. По большому счету, он согласен с Куайном, что идеальным объяснением было бы физическое объяснение (в его терминологии, на основе физической установки). Однако получить его нереально, поскольку между поведенческими паттернами и физическими процессами нет тождества.

Как мы уже сказали, «изюминкой» метафизики Деннета является объяснение сознания на основе эволюционизма. По своему типу его физикалистский монизм скорее является органицизмом или процессуальной философией (ее еще называют «философией Allife»). В отличие от других процесс-философов он толкует *эволюционизм как абстрактный принцип* и даже метафизический или космологический принцип. Он не ограничивает его действие специфически биологической территорией и ее законами, распространяя на все сферы бытия — природу, язык, мораль, искусство, религию, одним словом, на все артефакты культуры, включая искусственный интеллект. Одна из важных посылок метафизики Деннета состоит в том, что эволюционные процессы *алгоритмичны*. Алгоритмы работают объективно, автономно, нейтрально и независимо от каких-либо «целей», «смыслов» или «космического интеллекта». Их действие проявляется в разнообразии, репликации, имитации, случайностях, инновациях, пробах, ошибках, тестировании, отборе, адаптации, информации, коммуникации и многом другом. Все это видимое и невидимое многообразие является продуктом анонимных алгоритмических процессов. Они сочетают в себе случайность и отсутствие порядка с действиями естественного отбора и адаптации, производящих непрерывную ревизию сложившимся структурам. Процессуальный принцип эволюции — «Исследование и Развитие» — присущ растениям, животным, человеку, социальным системам, компьютеру.

Не только ответ на вопрос «как возможно сознание?», но и ответ на вопрос «как возможен компьютер?» Деннет пред-

388

лагает искать в алгоритмичности эволюционных процессов. Аналогия между функционированием сознания и функционированием компьютера у него не ограничивается констатацией сходства их когнитивных операций: метафизическая или космологическая сторона этого сходства состоит в том, что и сознание, в котором мысли сами себя мыслят, и компьютер, выдающий новую информацию на основе имеющейся информации, возможны потому, что существуют глобальные



алгоритмичные процессы, проявляющиеся и в человеческом сознании, и в искусственном интеллекте, и в артефактах культуры. Открытие «Универсальной машины Тьюринга» и создание компьютеров в принципе было бы невозможно без объективного, независимого от какого-то агента, протекания этих процессов. В конечном счете, все они являются реализациями проектов Главного Дизайнера — Матери-Природы.

Сориентированный на эволюционизм как абстрактный (универсальный) принцип, верящий в адаптивные и инновационные процессы в производстве сознания, Деннет видит путь к *нейтрализму* в отношении физического и ментального, прежде всего, в функционализме «логических состояний». До пришествия компьютеров философы сопрягали сознание либо с состояниями души, либо с состояниями мозговых процессов. С появлением компьютеров стало возможным сопрягать сознание с «логическими состояниями» машины Тьюринга и «отношениями логических состояний». В создании искусственного интеллекта он увидел инструмент, позволяющий по-новому посмотреть на возможности достижения нейтрализма, на перспективу преодолеть редукционизм и снять оппозицию физикализма и ментализма. Подобно тому как построенная на имитации и репликации работа компьютерной программы описывается на нейтральном языке, точно так же функциональные проявления сознания, построенные на имитации и репликации логических состояний, не нуждаются ни в специфически физическом, ни в специфически ментальном языке. Для их описания лучше всего подходит язык «Машины Тьюринга».

389

У Рассела и раннего Витгенштейна единство мира достигалось на основе изоморфизма логической структуры языка логической структуре мира. Деннет скептически относится к такому изоморфизму, как и к психофизическому параллелизму. Возможность обоснования единства мира у Деннета достигается на основе изоморфизма функциональных, логических, информационных процессов сознания и информационных процессов мира. Иначе говоря, изоморфизм между миром и сознанием осуществляется не на уровне физической или логической структур, а на *процессуально-информационном уровне*.

В эволюционистской метафизике Деннета сознание лишается «постоянной прописки» на площади уникального внутреннего мира субъекта и получает, так сказать, «планетарное гражданство», растворяясь в функционировании различного рода «машин», работающих в природе, в языке, социальности, культуре и в находящейся «нигде» Информационной Диаспоре. В своей гипотезе Деннет не заходит так далеко, чтобы говорить о «глобальном сознании». Однако не отмечает идею, что благодаря языку и другим семиотическим системам *сознание отдельного человека встроено в такую мировую систему коммуникации и информации*, которая, включая его в свои сети, дает возможность обрести индивиду свой неповторимый «центр нарративной гравитации». Более того, он предлагает посмотреть на homo sapiens как на «нейроны Земли». Сотворив язык и культуру, говорит он, «планета в итоге создала ее собственную нервную систему: нас»<sup>1</sup>.

Говоря о философии Деннета в контексте современной философии, нельзя обойти тему *сциентизма*. В нашей литературе сциентизм обычно ассоциируется с интенцией превратить философию в отрасль научного знания. У Деннета нет такой интенции. Чтобы яснее представить специфику его взглядов, сопоставим их с неопозитивистским вариантом сциентизма, который больше известен в нашей стране.

<sup>1</sup> Dennett D. We Earth Neurons // From Internet: <http://ase.tufts.edu/cogstud/pubpage.htm>. P. 2.

390

Деннет редко ссылается на Р. Карнапа, Г. Рейхенбаха, М. Шлика, тем не менее он очень многое вобрал из их идейного наследия (эхо их идей чувствуется во всей аналитической философии). Это сказывается прежде всего в уважении к науке и научной рациональности; эпистемический авторитет науки для него непререкаем. Гетерофеноменологический метод по сути дела является методом интерсубъективной проверяемости, за который ратовали неопозитивисты. Роднит его с неопозитивизмом убеждение, что развитая философия должна одновременно и «подпитываться» наукой, и ориентировать науку: наука может усилить свои эвристические возможности, если будет

применять философский инструментарий в анализе ее концептуальных оснований. В частности, он убежден, что предложенная им концепция сознания может помочь преодолеть сегодняшние тупики когнитивных наук. Одну из своих задач он видит в том, чтобы подготовить почву для превращения философии сознания в отрасль философии науки — респектабельный предмет рационального исследования, имеющий прикладную значимость. Одним словом, в согласии с неопозитивистами (и в разногласии с Рорти) Деннет убежден, что философия, при всех ее особенностях, может и должна иметь *теоретический статус*. Это значит, что она способна стать доказательной, аргументативной, подтверждаемой деятельностью сродни той, которая практикуется в науках. Отсюда его категорическое неприятие постмодернизма и релятивистского историцизма.

И все же сциентизм Деннета по своему типу иной, нежели у неопозитивистов. Его специфика во многом определяется, во-первых, тем, что, следуя традиции Витгенштейна и Райла, он работает с естественным языком, считая формальный (идеальный) язык, на который ориентировались неопозитивисты, ограниченным. Во-вторых, он ориентируется на другой образ науки. Если неопозитивисты отталкивались от осмысления физики, математики и логики, что определило жесткость их физикализма (и принципов рациональности), Деннет обратил свой взор на эволюционную биологию, нейронауку, психоло-

391

гию, теории искусственного интеллекта. Его физикализм, скорее, является и *органицизмом*, а защищаемая им рациональность является *«свободно плавающей рациональностью»* или *эволюционным когнитивизмом*.

Посылка о единстве Вселенной и тезис о невозможности объяснения сознания на основе ее раскола являются общими и для неопозитивистов, и для Деннета. Однако первые были редуционистами, видевшими путь к объяснению сознания в трансляции языка психологии к языку науки (или в онаучивании бихевиористской психологии), Деннет же антиредуционист. У него более сложная задача: совместить детерминизм физикализма с целеполаганием, представить аргументы, доказывающие, что физикалистское (и механицистское) объяснение человека не элиминирует целеполагание (и такие понятия, как «личностное», «моральность», «ответственность»), а вполне с ним стыкуется. Насколько убедительны его аргументы, читатель может судить и по статье «Условия личностного» (в Приложении), и по материалу 7-й и 8-й глав нашей книги.

Разница типов сциентизма еще более заметна на уровне метафилософии. Вдохновлявшая не только Карнапа, Шлика и Рейхенбаха, но и Куайна идея низвести философию до вводной, спекулятивной части в магистральном движении к зрелой (идеальной) науке, Деннету представляется нереальной. Философия, убежден он, занимала и будет занимать срединное место между наукой и искусством, используя методы той и другого. С искусством ее сближает раскованность воображения и отсутствие жестких стандартов творчества. Свою собственную работу он сравнивает с работой скульптора, из глыбы камня высекающего возникающие в интуиции образы, или с нарративной деятельностью литературных критиков. Он не относит метафизику к дисциплине *non grata*. Философии не возбраняется прибегать к спекулятивным, в том числе метафизическим, гипотезам, выходящим за рамки имеющегося фактуального знания, соединять идеи, наработанные в разных дисциплинах, делать неординарные эвристические ходы. Одним из ее эффективных средств являются мысленные экспе-

392

рименты. Наука тоже пользуется ими, однако философия широко прибегает к образам, метафорам и другим средствам обыденного и художественного языков, что неуместно в науке. Одним словом, Деннет работает в интеллектуальном контексте другой эпохи, и она наложила печать на его сциентизм. В заключение нашего рассказа о философии Деннета имеет смысл сказать несколько слов о ее перспективности и влиятельности. Прогнозы в отношении прихотливой жизни философии, как показывает история, рискованны. Вспоминается, например, что в 1967 году Ричард Рорти в предисловии к антологии «Лингвистический поворот», подводя итоги развитию

логического эмпиризма и философии лингвистического анализа, а фактически ставя диагноз состоянию аналитической философии середины XX века, нарисовал шесть сценариев дальнейшего движения философии. Все они так или иначе были завязаны на «повороте к языку» и методологическом номинализме. Не исключался сценарий «конца философии» как теоретической деятельности и переход к «постфилософской культуре». В своих «сценариях» Рорти явно недооценил возможность мощного «натуралистического поворота» в недрах лингвистической философии с фокусировкой извечных (картезианских и локковских) проблем, а тем более возможность интенсивного вовлечения философов в компьютерные дела. Учитывая рискованность любых прогнозов, нам все же кажется, что философия Деннета имеет шансы на развитие, модификацию, создание определенной традиции. Это мнение основывается на наблюдении, что тема натурализма слишком прочно вошла в американскую философию (и в американскую культуру, где ее даже называют «дарвиновской культурой»), чтобы быть вытесненной в ближайшие годы чем-то новомодным. Тема стыковки социального и биологического является актуальной не только для философии сознания, но практически для всех наук, занимающихся человеком. Вероятнее всего, перспективными окажутся те философские теории, которые представят более или менее правдоподобные гипотезы относительно механизмов соединения довербального и вербально-

393

го, когнитивного и феноменального. Оригинальная, во многом провокативная, но насыщенная неординарными гипотезами концепция Деннета вряд ли будет отодвинута на обочину этого магистрального движения.

Теперь о влиятельности философии Деннета. Судя по литературе, у него не так много адептов (есть несколько «деннетистов»). На самом деле его влияние огромно. Его имя постоянно фигурирует в статьях авторов, печатающихся не только в философских, но и в специальных журналах, таких как «Journal of Consciousness Studies», «Consciousness and Cognition», в психологических журналах, в периодике когнитивных наук. Не говоря уже о том, что авторы новых книг по проблеме сознания так или иначе высказывают свое отношение к идеям Деннета. Если говорить о содержательном влиянии, можно отметить, что у критической и конструктивной частей философии Деннета разные судьбы. Его атака на феноменализм и ментализм многими была встречена с пониманием и одобрением. Известные психологи Тони Марсель, Бернанд Баарс, нейрочеловек Марсель Кингсборн и многие другие профессиональные ученые в своих исследованиях широко прибегают к критическим аргументам Деннета против традиционного эмпиризма. Это и понятно: в целом его атака соответствует духу традиции англоязычной мысли, сложившейся под влиянием позднего Л. Витгенштейна, Дж. Райла, Дж. Остина, У. Селларса и других антифеноменалистов. У Модели Множественных Набросков другая судьба. Многие обескуражены ее необычностью, а также тем, что в ней осталось немало неясного. Значительно более позитивно встречена теория интенциональности: ученые говорят, что предложенный Деннетом метод объяснения поведения на основе интенциональной установки дает впечатляющие результаты во многих областях, где нужно понять резоны людей, когда они предпринимают какие-либо действия. Этот метод работает и при объяснении поведения животных.

С нашей точки зрения, то есть с позиции историка, интересующегося динамикой и пролиферацией идей, одна из самых главных заслуг Деннета состоит в том, что своей концепцией

394

он основательно подлил масла в огонь дебатов о сознании и стимулировал многих философов определиться в дилемме «монизм или дуализм». Увеличившийся поток литературы последних двух десятилетий свидетельствует о том, что не только философы, но и ученые взялись за перо и ринулись искать новые идеи, аргументы, придумывать мысленные эксперименты, пытаясь найти ответ на извечный вопрос: «Что такое сознание?» Короче говоря, ценность его вклада не зависит от того, покончил ли он с «головоломками» и представил ли он сознание «объясненным». В конце концов, роль философии в культуре состоит в том, чтобы проигрывать различные варианты гипотез, которые подготавливают почву для будущих решений. И чем больше альтернатив

предлагается в ее рамках и более ожесточенными являются баталии, тем больше шансов найти искомое решение.

Особо следует сказать о влиянии его стилистики философствования. В российской, да и в зарубежной литературе аналитическая философия часто предстает в образе «технической философии», занимающейся узкими проблемами, понятными для небольшого круга профессионалов. Ранние работы Деннета соответствовали этому образу, в поздних появились элементы художественного стиля: они стали более популярными, с множеством зарисовок и метафор, диалогов, придуманных и реальных историй. Это настоящая философская проза. Раскованность его философского стиля, как нам кажется, повлияла на стилистику ведущихся сегодня дискуссий в философии сознания: они стали более открытыми, рассчитанными на более широкую аудиторию. Примечательная черта нынешней аналитической философии — давние битвы монистов и дуалистов, объективистов и субъективистов все больше выглядят как соревновательность в образах или даже как «война метафор».

Завершая наш рассказ о Деннете, повторим сказанное в начале книги. В нашей работе мы не ставили перед собой задачу проанализировать его взгляды с точки зрения их «правильности» или «неправильности». Историк у брать на себя

395

роль арбитра в таком серьезном вопросе, как природа сознания, несерьезно. Нужна другая квалификация. Чтобы быть более или менее объективными, мы постарались экспозицию взглядов Деннета давать в сопряжении с экспозицией взглядов оппонентов. Однако и в этом случае судить, кто из спорящих ближе, а кто дальше стоит от истины, мы воздержались. (Предоставили возможность читателю самому быть судьей. Правда, для этого ему потребуется проделать теоретико-проблемный анализ, т.к. историко-философская дискрипция в лучшем случае может только обрисовать контекст.) Наша цель была предельно простой — как можно адекватнее изложить философские взгляды и интенции Деннета. Ее выполнение оказалось в высшей степени сложным и трудоемким делом. И не только в силу известных трудностей при изложении аналитических теорий, к тому же оперирующих научной информацией. Главная трудность состояла в том, что материалом выступала живая мысль творчески активного философа. В случае исторической реконструкции мысли прошлого мы воспринимаем ее во временной законченности и определенности. К тому же в распоряжении историка есть труды его коллег по профессии, на которые он может опираться. В случае исследования творчества здравствующего философа ты один и идешь по целине. Время интеллектуального героя еще не истекло, система взглядов не обрела завершенности; он продолжает пояснять, отвечать на критику, корректировать и дополнять ранее сказанное и т.д. От него ожидают новых идей. Нам кажется, есть основания предполагать, что Деннет напишет книгу, где он более определенно изложит метафизические воззрения. Заведомо можно сказать, что в ней будет много неординарных мыслей.

Я хочу подчеркнуть, что эта книга обращена к молодому читателю, которого притягивает загадочность феномена сознания и чья любознательность не мирится с тем, что отгадка еще не найдена. Мы надеемся, что представленный в книге материал сыграет роль вспаханной целины, на которой ему легче будет работать над разгадкой тайны сознания.

## Приложение 1 Тексты Д.К. Деннета (Перевод с английского Н.С. Юлиной)

### 1. По следу Дарвина. Где есмь Я?

Президентский доклад на заседании Американской философской ассоциации. Декабрь 2000 года.\* На этой неделе философия потеряла гиганта. На Рождество в возрасте 92 лет умер Уиллард Ван Орман Куайн. Нет необходимости напоминать коллегам по Американской философской ассоциации о его огромном вкладе в философию. Пользуясь моментом, я хотел бы сказать, что он значил для меня лично. Он был не только моим учителем и другом; он был моим героем, человеком, который спокойно и вместе с тем страстно был предан истине, ясности, пониманию всего, что есть под солнцем, — и старался, чтобы сам он был понят. Больше чем кто-либо другой он заставил меня гордиться тем, что я являюсь философом. Поэтому я хотел бы посвятить мое Президентское Послание его памяти.

\* Dennett D.C. In Darwin's Wake, Where am I? Presidential Address. American Philosophical Association, December 29, 2000. Internet: <http://ase.rufts.edu/cogstud/pubpage.htm>. P. 1.  
397

\*\*\*

Parfois je pense; et parfois, je suis.  
*Paul Valéry*<sup>1</sup>

Поль Валери в его «*Variation sur Descartes*» нарисовал превосходный образ акта исчезновения, который стал преследовать философию с тех пор, как Дарвин опрокинул картезианскую традицию. Если *мое тело* представляет собой не что иное, как команду из нескольких триллионов бессознательно взаимодействующих между собой роботоподобных клеток для производства всех крупномасштабных паттернов, которые традиционно считались продуктами деятельности моего сознания, — не являющейся механической, — получается, что не остается ничего, что было бы *моим Я*. В тени Дарвина скрывается страшилка, невероятная вещь — Исчезающая Самость<sup>2</sup>. Один из самых ранних критиков Дарвина, предвидя грядущее и едва сдерживая возмущение, писал: В теории, с которой нам придется иметь дело, роль творца передана Абсолютному Невежеству; поэтому в качестве фундаментального принципа данной системы в целом можно было бы представить следующее: ДЛЯ СОЗДАНИЯ СОВЕРШЕННОЙ И КРАСИВОЙ МАШИНЫ НЕТ НЕОБХОДИМОСТИ ЗНАТЬ, КАК ЕЕ ДЕЛАТЬ. Если внимательно вдуматься в него, этот принцип в сжатом виде выражает существенный смысл Теории, рисуя в нескольких словах все то, что проделал м-р Дарвин; прибегнув к странной инверсии рассуждения, он, по-видимому, счел Абсолютное Невежество полностью квалифицированным для занятия места Абсолютной Мудрости во всех достижениях творческого таланта<sup>3</sup>.

<sup>1</sup> Порой я мыслю, порой существую. (Фр.) Поль Валери. [Valéry P. Cahiers. Ed. by J. Robinson, 2 vols. (Paris: Edition de la Pleiade 1973–74). V. 2 (1974). P. 1388.

<sup>2</sup> Dennett D. Elbow Room: the Varieties of Free Will Worth Wanting. Oxford, N.Y., 1984. P. 13.

<sup>3</sup> MacKenzie R.B. The Darwinian Theory of the Transmutation of Species Examined. 1868 (опубликовано анонимно «By a Graduate of the University of Cambridge»). London: Nisbet & Co. Quoted in a review, Athenaeum. № 2102. Febr. 8. P. 217.

398

Эта «странная инверсия рассуждения» обещает (или угрожает) аннулировать источник творческого потенциала — картезианский *res cogitans*. Где же тогда окажемся мы? По-видимому, Нигде. Кажется, что при «сведении» творчества до уровня «простого механизма», нетрудно будет показать, что *мы* как таковые вроде бы не существуем вообще. Или существовать-то мы будем, но не в качестве мыслителей, проявляющих подлинную «Мудрость во всех проявлениях творческого таланта». Статус Индивида как Автора и Свершителя дел будет понижен: кажется, что в этом случае человека можно представить в виде простого узелка в фабрике каузальности.



Всякий раз, погружаясь в акт творчества, у нас появляется чувство, что сам он теряется из виду. Гениальность, которая, как кажется, видна на расстоянии, в конечном итоге подменяется глупой машиной, эхом шокирующей дарвиновской подмены Абсолютной Мудрости Абсолютным Невежеством в процессе творчества биосферы. Многие люди всем своим нутром не приемлют дарвинизм. При всей смутности и темности их антипатий, этот настрой всегда поражал меня своей глубинностью в смысле наибольшей стойкости и наименьшей доступности рациональной критике. Есть немало мыслящих людей, с иронией относящихся к креационизму, с порога отбрасывающих дуализм, защищающих академический гуманизм, но стоит им предложить основывающуюся на дарвинизме концепцию творческого интеллекта, они тотчас же начинают ежиться. Идея о том, что все проявления человеческого гения могут быть *в конечном итоге* поняты в виде механистически произведенных продуктов каскада алгоритмов — генерируй-и-тестируй, у этих людей, повторяю, глубоко, непредубежденно и открыто мыслящих в отношении других предметов, вызывает внутреннюю неприязнь.

Абсолютное Невежество? Чур всех тех, кто хочет соединить «Я» и «AI»!<sup>1</sup> Интуитивная прозорливость — это источник эво-

<sup>1</sup> «AI» — искусственный интеллект, в дальнейшем этот термин мы даем в принятой английской аббревиатуре.

399

люции, поэтому достаточно, чтобы эволюционист, каким являюсь я, должным образом приспособил удачно выделенную крупным шрифтом фразу МакКензи для достижения цели, которую он едва ли мог помыслить. Спустя более чем столетие нескрываемый сарказм МакКензи по поводу передачи Дарвином полномочий Абсолютной Мудрости Абсолютному Невежеству удивительно созвучным эхом отозвался в таком же сарказме в отношении людей, уверовавших в тезис, который Джон Сёрль<sup>1</sup> назвал «сильным AI». Иначе говоря, людей, считающих, что *реальный* интеллект может быть создан искусственно, что различие между наделенными и лишенными сознания механизмами есть различие в дизайне (или *программе*), поскольку все, что можно спроектировать в hardware, можно имплементировать в виртуальной машине, имеющей такую же компетенцию<sup>2</sup>. Дарвиновская «странная инверсия рассуждения» переворачивает старое представление о творчестве с головы на ноги. Согласно прежней, или «нисходящей» (top-down), перспективе творческого интеллекта, предполагается, что для создания меньшей вещи всегда требуется какая-то большая, умная и наделенная воображением вещь. Никакая подкова никогда не

<sup>1</sup> Searle J. Minds, Brains and Programs // Behavioral and Brain Sciences. V. 3. 1980. P. 417—58.

<sup>2</sup> Это с очевидностью истинно в отношении всех компетенций процессов информации или контроля, но не в отношении продуктивных или трансформационных процессов вроде лактации, требующих носителя и совокупности специфических материалов. Поскольку целью Сёрля является различие «контролирующих сил» мозга и «восходящих снизу каузальных сил», «продуцирующих интенциональность», некоторые комментаторы думают, что он считает интенциональность чем-то вроде особой, выделяемой мозгом субстанции. Поскольку сам он отрицает такую трактовку, это обязывает его представить какой-то другой способ различения этих мистических каузальных сил и контролирующих сил, которые могут быть имплементированы в software, и объяснить нам, почему они не могут быть имплементированы в виртуальной машине.

400

выковала кузнеца; никакой горшок не вылепил гончара. Следовательно, и мы сами, и все окружающие нас необыкновенные вещи должны быть созданы кем-то, наделенным еще большим воображением; кем-то вроде нас, но только еще в большей мере. Для многих, возможно, для большинства людей эта идея представляется *совершенно очевидной*. Вот что говорится в полученном по почте пропагандистском листке креационистов:

1. Знаете ли вы какое-либо строение, у которого не было строителя? ДА / НЕТ.
2. Знаете ли вы какую-либо картину, у которой не было художника? ДА / НЕТ.
3. Знаете ли вы какой-либо автомобиль, который не имел изготовителя? ДА / НЕТ.

Если вы ответили «ДА» на любой из вышеупомянутых вопросов, приведите детали.

Предложенный Джоном Сёрлем мысленный эксперимент, известный как «Китайская Комната», является вариантом этой безысходной шутки креационистов:

Знаете ли вы любую машину, которая может понимать китайца? ДА / НЕТ.

Если вы ответили «ДА», приведите детали.

Как бы сильно ни импонировала эта идея здравому смыслу, Дарвин убедительно показал, каким образом она может быть ложной. Он показал, что в действительности не только можно вообразить теорию, строящую объяснения «снизу-вверх» (bottom-up), но и доказать ее очевидную правоту опытным путем. Доказать, что Абсолютное Невежество *является* в полной мере квалифицированным для занятия места Абсолютной Мудрости во всех достижениях творческого мастерства — *любого* мастерства.

Риторические вопросы креационистов — это только жесты в сторону предполагаемых трудностей, возникающих перед любым человеком, осмелившимся «привести детали» примера творения «снизу-вверх». Вызов же Сёрля представляет собой обзор возможных ходов, которые могли бы предпринять сто-

401

ронники «сильного AI» в попытках «сообщить детали», а его конечная цель — опровергнуть все эти ходы до одного. Опровергающие старания Сёрля мало задевают сторонников сильного AI, а сопоставление его позиции с позициями креационистов показывает, почему это так. Биологи, которым *еще не удалось* объяснить какую-либо отдельную загадку на вовсе не таинственном пути, ведущем от одного чуда природы к другому, и которые *еще не* могут «поведать о деталях», удовлетворяющих вездливому критика, тем не менее имеют весьма внушительный перечень достижений в предоставлении деталей, а также стабильную и плодотворную теорию, применяемую в приведении и подтверждении новых деталей. Поэтому они просто отмахиваются от риторического вывода: «Вы никогда не добьетесь успеха!» Они спокойно осознают, что им нужно добавить еще кой-какие новые штрихи, прежде чем заявлять об успехе.

Как и биологи, сторонники сильного AI не отрицают, что все созданные на сегодня модели AI во многих отношениях несовершенны, порядок их величины слишком прост, к тому же многие из них отстаивают особый взгляд на AI, который просто ошибочен. Вместе с тем они обращают внимание на то, что Сёрль бросил критический вызов не отдельным деталям предпринимавшихся доселе попыток; в его намерение входит формулировка аргумента, раз и навсегда доказывающего *принципиальную* невозможность сильного AI. Защищаемый им вывод нацелен охватить все *возможные* сложности их базовых теоретических рамок. В противовес Сёрлю, сторонники AI исходят из того, что теоретический каркас, на основе которого строится их деятельность, есть не что иное, как прямолинейная экстенсия и в человеческий мозг, и во все его периферийные дивайсы и интерфейсы дарвиновской программы механизмов, не имеющих сознания и тем не менее в конечном итоге проделывающих всю работу. Если дарвиновские механизмы могут объяснить существование жаворонка во всей его изумительности, они, конечно, могут объяснить и существование оды

402

в честь соловья<sup>1</sup>. Поэма — замечательная вещь, хотя на самом деле не более удивительная, чем живой, поющий жаворонок.

Живучесть безосновательных антипатий часто объясняется их сокрытием под другими личинами: они подмешиваются в правомерные возражения против преувеличений, присущих теории, находящейся под огнем атак. Так как энтузиазм дарвинистов всегда превышал возможности охвата многообразия явлений с помощью дарвиновской теории, всегда появлялась трезвая критика его эксцессов. Сказанное относится и к эксцессам идеологов AI. По мере того как накаляется сражение, возникают как подозрение, так и понимание. Дарвинисты, — а я являюсь одним из них, — уверены в том, что софистичный дарвинизм с его тонкими деталями является доказательством, опровергающим все возражения и недоверие. Хотя дарвинистам не следует забывать судьбу

<sup>1</sup> Эта перспектива помогает нам понять сердечную приверженность многих сторонних наблюдателей к различным, кажущимся им *очевидными* и давно имеющими хождение

альтернативам дарвиновскому механизму. В последние годы самыми известными из них были: апелляция к «самоорганизации» «на краю хаоса» (Stuart Kauffman, Per Bak и др.), «динамичная системная теория», охватывающая эволюцию и познание (Esther Thelen, Walter Freeman, Timothy van Gelder и др.), и, конечно, настойчивые заявления Стивена Гоулда, что эволюция не является фундаментально алгоритмическим процессом. Гоулд воинственно настроен против моей позиции, опирающейся на результаты работ таких авторов, как Дарвин, Фишер и Холдейн, а также Уильямс и Мейнард Смит. Когда рассеется дым баталей, идеи критиков дарвинизма, в лучшем случае, будут расценены как интересные усложнения базисных дарвиновских механизмов, точно так же, как коннекционистские архитектуры и модели воплощенного познания являются интересными усложнениями базисных идей AI. Все эти столкновения позиций, в лучшем случае, являются конструктивными разногласиями по поводу «приведения деталей», а не вызовом базисному дарвиновскому видению.

403

раздражающих претензий фрейдистов 50-х и 60-х годов, требовавших рассмотрения всего и вся через призму категорий их героя и в итоге очутившихся перед фактом, что, приспособившая их *ко всему на свете*, они разжижали фрейдизм, в результате чего в значительной части он оказался Пиквикским Фрейдизмом.

Иногда сигара — это только сигара, и иногда идея является просто идеей — не меме; а иногда какой-то кусочек машинной ментальности не вполне разумно интерпретируется как адаптация, истоки которой восходят к временам нашего предка охотника-собирателя или даже к еще более ранним временам. Хотя очевидно, что этот кусочек произошел (с модификациями) в результате некоторой комбинации или каких-либо других видов адаптаций. Будучи дарвинистами, мы стараемся напоминать себе об этом, надеясь, что наши упорные противники придут к осознанию, что дарвиновская теория творческого потенциала содержит в себе не только обещающее решение, она является единственным решением проблемы, которая является проблемой каждого из нас: *как может совокупность сотен миллиардов не имеющих сознания нейронов создать творческий ум, мое Я?*

Брошенный дарвинизму критический вызов кратко сформулировал Уильям Паундстон в известной фразе о «старой фантазии об обезьяне, случайно напечатавшей "Гамлета"». Он вычислил, что вероятность такого случая «1 из 50 помноженный на 0,000». В свете его вычислений вообще кажется поразительным само существование чего-то столь сложного, как текст Гамлета. Заявление, что «Гамлет» был написан Шекспиром, а не некоей случайной силой, только переворачивает проблему. Шекспир, как и все остальное в мире, должен возникнуть (в конечном счете) из гомогенной и существовавшей до него Вселенной. Как бы ни относиться к этой идее, «Гамлет» — продукт первичного хаоса<sup>1</sup>.

<sup>1</sup> Poundstone W. The Recursive Universe: Cosmic Complexity and the Limits of Scientific Knowledge. N.Y., 1985. P. 23.

404

Но откуда взялся весь этот дизайн? Какие вообще можно помыслить процессы, способные породить столь невероятные «достижения творческого искусства»? Дарвину виделось следующее: дизайн всегда является не только ценным, но и дорогостоящим приобретением. Это не падающая с небес манна, в нем должен быть аккумулирован трудный путь, потребовавший много времени и энергоемких процессов для осуществления бессознательного поиска через пространство «первичного хаоса» и для одновременного автоматического сохранения имевших место счастливых случаев. Не имеющий жестких креплений процесс Поиска и Развития (Research and Development, далее R&D — *прим. пер.*) удивительно неэффективен. Однако — и в этом состоит величайшее прозрение Дарвина — если дорогостоящие плоды R&D бережливо накапливаются, копируются и подвергаются многократному использованию, они могут быть сохранены, а со временем вылиться в «достижения творческого искусства». «Этот принцип сохранения, — говорит Дарвин, — я назвал, ради краткости, естественным отбором»<sup>1</sup>.

В том, каким естественный отбор виделся Дарвину, не содержалось никакого условия, чтобы процессы R&D происходили всюду и всегда, в одном и том же темпе, с той же самой эффективностью или неэффективностью. Взгляните на поразительно огромное и многомерное

пространство всех *возможных* продуктов этого дизайна — как естественных, так и искусственных. Каждый мыслимый кит и единорог, каждый автомобиль, космический корабль или робот, каждая поэма, математическое доказательство или симфония находят свое место где-нибудь в этом Пространстве Дизайна (Design Space). Если представить работу дизайна — или R&D — в виде своего рода *подъема* (lifting) в Пространство Дизайна<sup>2</sup>, можно уви-

<sup>1</sup> Darwin Ch. Origin of Species. Ch. 4 summary (P. 127 in facs. edition).

<sup>2</sup> Тактика пространственного изображения эволюционных процессов и их результатов является естественной и часто используемой метафорой, применяемой в моделях восхождения в горы и особенно

405

деть, что постепенный, часто откатывающийся назад, максимально неэффективный базисный поисковый процесс в каких-то важных случаях способен порождать новые условия, ускоряющие весь этот процесс и позволяющие быстрее и эффективнее осуществлять локальный подъем<sup>1</sup>. Назовите любой такой продукт раннего R&D *подъемником* или подъемным устройством (stane) и сопоставьте с тем, чего, согласно дарвинизму, не бывает: наития свыше, *сошедшего с небес озарения* (skyhook — небесных зацепок)<sup>2</sup>. Подобно манне с небес, наитие свыше было бы чудом, и если какой-либо автор прибегает к нему в «объяснении» творчества, значит, он фактически подписался под своим поражением. Его символом является фраза: «Здесь имеет место чудо»<sup>3</sup>.

популярна в адаптивных ландшафтах (упомяная наиболее очевидные и популярные ее применения). Разумеется, естественность метафоры не гарантирует ее здравость и даже может маскировать ее ограниченность, но поскольку было доказано, что базисная стратегия картографии проявила себя особенно плодотворной *в критике* свехупрощенных эволюционных идей (e.g., Kauffman's «rugged landscape», Eigen's «quasi-species»), ее не стоит рассматривать как очевидное отклонение в сторону упрощенного видения дарвинизма.

<sup>1</sup> Smith M.J., Szathmary E. The Major Transitions in Evolution, Oxford, 1995. Авторы идентифицируют девять случаев (больших транзаций) интенсификации эволюционных процессов, создающих своего рода подъемники.

<sup>2</sup> Dennett D. The Tools for R and D: Skyhooks or Cranes? // Dennett D. Darwin's Dangerous Idea. Evolution and the Meaning of Life. L., 1995. P. 73—80.

<sup>3</sup> Имеется в виду знаменитая карикатура Сиднея Хэрриса, на которой изображена висющаяся в физической аудитории доска, испещренная внушительными формулами, в середине которых в скобках написана фраза: «Я думаю, что на втором шаге вам следовало быть более ясным». (Воспроизведена в: Dennett D. Consciousness Explained, 1991. P. 38.)

406

Ну хорошо, чем же тогда является сознание? Ответ дарвиниста прямолинеен. Сознание — это подъемник, созданный из подъемников и создающий подъемники. Оно представляет собой некий механизм огромной (но не невообразимой) сложности, способный продираться через Пространство Дизайна головокружительными (но не чудесными) шагами, благодаря всем предшествующим R&D и всем источникам, подвергнутым им эксплуатации. Как реагирует на это антидарвинист? Самым превосходным образом его реакцию выразил Пикассо, один из творческих гениев XX столетия (хотя, как и МакКензи, он, скорее всего, не вкладывал в свои слова смысл, который я намерен вложить в них): Je ne cherche pas; je trouve (фр.: я не ищу; я нахожу).

Пикассо хотел подать себя публике в образе настоящего гения, человека, которому нет нужды участвовать в черной работе проб и ошибок, в производстве-и-тестировании, в R&D. Хотел представить себя способным сразу прыгнуть на пик вершины — прекрасные дизайны — в обширном Пространстве Дизайна, не имея надобности прокладывать траекторию (или заниматься ее поисками) побочными испытаниями на любом отрезке пути. Как вдохновенная доля хвастовства, это *non pareil* (фр. — претензия на исключительность), но я не верю ему ни на секунду. Любой человек, кому случилось пройти по выставке рисунков Пикассо (что посчастливилось мне недавно в Валенсии) и воочию увидеть множество вариаций на одну и ту же тему, причем все они подписаны — и проданы — художником, может убедиться, что это не так. Безотносительно к тому, что может означать это *bon mot* (фр. — красное словцо) Пикассо, он не



был искренним, когда утверждал, что не занимался энергоемким, отнимающим много времени исследованием окрестностей Пространства Дизайна. В лучшем случае он мог бы сказать, что его собственные исследования были столь профессионально наработаны и настолько эффективны, что самому ему все это вовсе не казалось *работой* над дизайном. Тем не менее остается вопрос, что же особенное было внутри Пикассо, что сделало его таким хорошим дизай-

407 нером? Наитие с небес или превосходный набор подъемников?<sup>1</sup>

Здесь уместно перейти к характеристике взаимной подозрительности, сложившейся между дарвинистами и антидарвинистами, мешающей эмпирическим исследованиям феномена творчества. Дарвинисты уличают противников в апелляции к наитию свыше, видят у них стремление найти удивительный дар гения, чья мощь неразложима на механические акты, какими бы сложными и насыщенными информацией ни были предшествующие ему процессы R&D. Антидарвинисты подозревают своих противников в склонности к объяснению творческих процессов, принижающих роль Искателя, Автора, Творца: эта роль исчезает или, в лучшем случае, превращается в простую временную траекторию процесса бессознательной репликации. Я полагаю, что небольшого прогресса в споре можно достичь, если на примере об обезьяне, приведенном Паундстоном, представить построение *творения творца Гамлета*. Для этого проведем небольшой мысленный эксперимент.

Предположим, что доктор Франкенштейн создал дизайн и сконструировал монстра — «Спейкшира», который затем садится за стол и пишет пьесу под названием «Спамлет». Мой вопрос относится не к автору *Веверлея*, а к автору «Спамлета».

Кто является автором «Спамлета»?

Сначала оговорю моменты, которые я считаю несущественными для проведения данного мысленного эксперимента. Я не сказал, является ли «Спейкшир» роботом, сконструированным из металлических и силиконовых чипов или, как это было с монстром Франкенштейна, созданным из человеческой ткани, или клеток, или протеинов, или аминокислот, или углероди-

<sup>1</sup> Мне не удалось найти источник этого заявления Пикассо, но его слова хорошо сбалансированы известной ремаркой более трезво мыслящего творческого гения — Томаса Эдисона: «Гений — это один процент вдохновения и девяносто девять процентов потения». Интервью, опубликованное в: Life [1932]. Ch. 24. [Oxford Dictionary of Quotations].

408

стых атомов. Пока идет дизайнерская работа и д-р Франкенштейн творит свою конструкцию, для нашего эксперимента не имеют значения материалы, из которых он ее создает. Вполне может оказаться так, что единственный путь сконструировать робот, достаточно маленький, достаточно шустрый и достаточно энергоэффективный (чтобы он мог сидеть на стуле и печатать пьесу), — создать его из искусственных ячеек, заполненных красиво обработанными моторными белками и другими основанными на углероде нанороботами (nanobots). Вопрос этот интересен с технической и научной точек зрения, но в данном случае он меня не интересует.

По этой же самой причине, если «Спейкшир» является роботом из металла и силикона, можно было бы допустить, что он больше, чем галактика, если требуется только внесение необходимых усложнений в его программу и снятие ограничений скорости света ради нашего мысленного эксперимента. Принято считать, что такого рода технические ограничения вполне допустимы в мысленных экспериментах, поэтому пускай это будет так. Если д-р Франкенштейн предпочтет сделать свой AI-робот из протеинов или из чего-нибудь вроде того, это — его дело. Если робот хорошо скрещивается с нормальными человеческими существами и, родив ребенка, приобрел способность к созданию нового вида существ, все это выглядит замечательно и вызывает игру воображения, однако в данном случае нас интересует другое — предполагаемое детище «Спейкшира», «Спамлет». Вернемся к нашему исходному вопросу: кто является автором «Спамлета»?

Чтобы схватить суть вопроса, нам следует заглянуть внутрь «Спейкшира» и посмотреть на то, что здесь происходит<sup>1</sup>. В одном крайнем случае мы обнаружим здесь файл (если «Спейкшир» — робот





из  $10^{84}$  кубических сантиметров (Eigen M. Steps Towards Life. Oxford, 1992. P.10).

411

либо интересное не бесконечно мал<sup>1</sup>. Слова «бесконечно малое», «бесконечно большое» привычным способом применяют к величинам, однако нам следует избегать их: к сожалению, все используемые для этих целей стандартные метафоры — «астрономически большое», «игла в стоге сена», «капля в океане» — выглядят до смешного ограниченными. Никакое *действительно* астрономическое количество (вроде количества элементарных частиц во Вселенной или времени, измеряемого в наносекундах, начиная с Большого Взрыва) даже не проглядывается на фоне этих огромных-но-конечных чисел.

Если бы том со смыслом в Библиотеке можно было найти столь же легко, как и искомую каплю в океане, мы были бы при деле! На самом деле, случайно оказавшись в Библиотеке, наш шанс натолкнуться когда-либо на том, содер-

<sup>1</sup> Весьма любопытно, что, несмотря на конечность Вавилонской I Библиотеки, в ее пределах содержатся все грамматические предложения английского языка. Они составляют бесконечное множество, в то время как библиотека конечна! Тем не менее, независимо от его длины, любое английское предложение может быть разбито на 500-страничные куски, каждый из которых находится где-то в библиотеке! Как это возможно? Некоторые книги могут использоваться более чем однажды. Понятнее всего расточительный случай: поскольку есть тома, в каждом из которых имеется единственный знак, а другие тома совсем пустые, повторное использование этой одной сотни томов может сотворить любой текст любой длины. У. Куайн в его весьма информативном и вызывающем восхищение эссе «Универсальная библиотека» (W.Quine. Quiddities: An Intermittently Philosophical Dictionary. Cambridge MA, 1987) заметил следующее: «Если прибегнуть к стратегии многократного использования томов и все перевести в код ASCII, который имеется в вашем Word-процессоре, вы сможете хранить всю Вавилонскую Библиотеку в двух очень тоненьких томах, в одном из которых напечатано 0, а в другом — 1! (Куайн также заметил, что фантазия универсальной библиотеки была предложена психологом Теодором Фехнером задолго до Борхеса).

412

жащий грамматическое предложение, настолько *Исчезающе мал* и настолько стремится к нулю, что нам стоило бы этот термин (Vanishingly Small) писать с прописной буквы и давать его в связке с термином *Наигромადнейший* (Vastly), кратким словом для обозначения Значительно-большо-чем-астрономическое<sup>1</sup>.

Занятно порассуждать о поразительной огромности этого конечного набора возможных книг по сравнению с любой фактической библиотекой. Большинство книг, как было замечено, является просто мусором. Поэтому рассмотрим исчезающее малое подмножество книг, составленных полностью из английских слов и не содержащих ни одного неправильного написания. Конечно, само по себе это огромное подмножество (Vast set) и содержащееся в нем, но с исчезающе малым шансом находимое, Огромное подмножество, в котором английские слова выстроены в грамматические предложения. В свою очередь, Огромное, но исчезающее подмножество этого подмножества представляет собой подмножество книг, составленных из английских предложений, реально имеющих смысл.

Их Огромное, но Исчезающее (стремящееся к нулю — *Н. Ю.*) подмножество относится к рассказу о ком-то — назовите его Джоном. А их Огромное, но Стремящееся к нулю подмножество — относительно смерти Джона Ф. Кеннеди. Из них Огромное, но стремящееся к нулю подмножество является истинным... А Огромное, но исчезающее подмножество возможных истинных книг относительно смерти Д.Ф. Кеннеди написано полностью в лимериках. В этом случае мы имеем многопорядковые величины более возможных подлинных книг в форме лимерика относительно смерти Д.Ф. Кеннеди, нежели книг, имеющих в Библиотеке Конгресса.

<sup>1</sup> У. Куайн (в цит. выше работе) для этих же целей использует термин «гиперастрономический». Два предыдущих параграфа взяты с небольшими изменениями из моей книги «Darwin's Dangerous Idea». P. 108—109.

413

Теперь мы подготовлены для возвращения к «игле в стоге сена» — к «Спамлету», чтобы

посмотреть, как его траектория к этому специфическому месту в Вавилонской Библиотеке проложила свой траверс в действительной истории. Если мы обнаружим, что все путешествие уже было закончено ко времени, когда память «Спейкшира» была выстроена и заполнена информацией, мы узнаем, что в данном поиске «Спейкшир» не играл вообще никакой роли. Если, оглядываясь назад, мы обнаружим, что единственная роль «Спейкшира» состояла в прогонке сохраненного текста через грамматический контроль, перед тем как использовать его в качестве руководства в осуществлении своих действий по печатанию, на нас не действуют заявления об авторстве «Спейкшира». Все это измеримая, но стремящаяся к нулю часть тотальности R&D.

Существует измеримая галактика почти близнецов-текстов «Спамлета» — приблизительно сотня миллионов различных уменьшенных мутантов, имеющих только один-единственный неисправленный типографский знак, и если расширить горизонт для включения этого единственного знака на страницу, мы начали бы входить в пространство Огромных чисел вариаций на эту тему. Двинувшись еще немного назад, как только мы поднялись от знаков к мысли (from typos to *thinkos*)<sup>1</sup>, к этим доказательно ошибочным или сверхоптимально выбранным словам, мы начнем входить в область серьезного авторства, отличающегося от простого редактирования копий. Относительная тривиальность редактирования копий, и в то же время его важность, которую нельзя игнорировать в формировании конечного продукта, получает хорошую репрезентацию в терминах нашей метафоры Пространства Дизайна. Здесь каждый маленький кусочек подъема что-то означает и временами устремлен к некоей совершенно

<sup>1</sup> Более подробное пояснение этого понятия см.: Dennett D. From Typo to Thinko: When Evolution Graduated to Semantic Norms. Материалы конференции о культурной эволюции (Fyssen).

414

новой траектории. Как обычно, на этом распутье мы можем процитировать Людвиг Мис ван дер Роэ: «Бог кроется в деталях».

По рекомендации Дугласа Хофстадтера теперь нам следует переключить кнопки в нашем мысленном эксперименте <sup>1</sup> и посмотреть на другую крайность: предположим, что д-р Франкенштейн передал большую часть работы «Спейкширу». Наиболее реалистичный сценарий был бы, конечно, такой: д-р Франкенштейн снабдил «Спейкшира» виртуальным прошлым или псевдовоспоминаниями об его опыте. И, реагируя на инсталлированное тем же д-ром Франкенштейном неудержимое желание писать пьесу, он получил возможность опираться на них в своей работе. Можно предположить, что среди его псевдовоспоминаний фигурируют проведенные в театре вечера, чтение книг, какие-то безответные влюбленности, шокирующие влечения, позорные предательства и т.п.

Что мы имеем сейчас? Возможно, что некоторые кусочки истории о «человеческом интересе», поданные в новостях средств массовой информации, послужат катализаторами и подстегнут «Спейкшира» вовлечься в неистовый процесс «генерирования-и-тестирования». Процесс будет состоять из поиска в его памяти полезных лакомых кусочков и тематики, из преобразований — перемещений или подавлений — того, что там находится, из утрясаний частей во временные, обнадеживающие структуры, находящиеся в конкурентной борьбе за завершенность. Большинство этих структур будут демонтированы коррозионными процессами внутренней критики, которая все же выделит полезные кусочки сейчас или потом, и т.д. И весь этот многоуровневый поиск каким-то образом управлялся бы изнутри многоуровневыми, внутренне генерируемыми оценками, включая оценку оценок... оценку оце-

<sup>1</sup> Douglas R. Hofstadter. Reflections (о Дж. Серле) // D. Hofstadter and D. Dennett. (eds.). The Mind's I: Fantasies and Reflections on Self and Soul. N.Y., 1981.

415

нивающих функций, представляющих собой реакцию на оценку... продуктов непрерывно идущих поисков<sup>1</sup>.

Если бы изумленный д-р Франкенштейн на самом деле мог предвидеть всю эту активность, начиная с ее мельчайшей частицы на максимально турбулентном и хаотичном уровне, и в его распоряжении имелось бы собственноручно созданное виртуальное прошлое «Спейкшира» со

всей его поисковой машиной для производства именно этого изделия — «Спамлета», тогда он еще раз был бы автором «Спамлета», но также, к слову сказать, и Богом. Такое Огромное предвидение было бы просто чудом. Дабы сохранить хоть какую-то долю реализма в нашей фантазии, можно установить кнопки в менее экстремальном положении. Предположить, что д-р Франкенштейн не был способен предвидеть все это в деталях и, скорее всего, делегировал «Спейкширу» большую часть трудной работы по завершению траектории в Пространстве Дизайна, *тому или иному литературному процессу*, или чему-то, что будет детерминировано последующими R&D, происходящими внутри самого «Спейкшира».

Простым нажатием кнопки мы сейчас прибыли в соседство с самой реальностью, ибо на самом деле у нас уже есть фактические примеры внушительных искусственных Авторов, в Огромной степени опережающих предвидение их собственных создателей. Сегодня еще не созданы достойные серьезного внимания драматические произведения, однако уже есть искусственный игрок в шахматы — созданный в IBM компьютер Deep Blue, и искусственный (компьютерный — *Н. Ю.*) композитор — ЕМІ Дэвида Коупа. Оба имеют результаты, *в каких-то отношениях* рав-

<sup>1</sup> Выражаю благодарность Тони Марселю за то, что он привлек мое внимание к этому пассажи. Сам Шекспир был, конечно, большим эксплуататором дизайнерской работы других и, подшучивая по поводу собственной репутации, цитировал критика, когда у того Autolycus описывал себя как «воришку непродуманных пусячков» (in *A Winter's Tale*, Act IV, scene III).

416

ноценные тем, которые достигнуты лучшими человеческими умами.

Кто победил Каспарова, чемпиона мира по шахматам? Победил его не Мюррей Кэмпбелл или кто-либо из его команды в IBM. Каспарова победил Deep Blue. Он придумал лучшую шахматную игру, нежели дизайны игры, придуманные для него кем-либо из создавшей его команды специалистов. Ни один из них не может считаться автором победной игры против Каспарова. А Deep Blue может. Все это так, но все-таки. Я уверен, что когда в шахматной игре Deep Blue одерживает победу над Каспаровым, у многих из вас наготове возражение, что грубые поисковые методы компьютера *совершенно* не похожи на исследовательские процессы, используемые Каспаровым в его прикидках возможных шахматных ходов. Однако это совсем не так — или, по меньшей мере, это не так в единственном значении, которое могло бы иметь смысл в контексте наших дебатов об универсальности дарвиновской перспективы применительно к творчеству. Мозг Каспарова состоит из органических веществ, и его архитектура в каких-то весьма важных аспектах непохожа на архитектуру Deep Blue. Тем не менее его мозг, насколько мне известно, представляет собой огромную параллельно работающую машину, в течение длительного времени выстроившую превосходный вектор экономной поисковой техники, помогающий ему не тратить время на ненужные ходы. Нет сомнения, что инвестиция в R&D в этих двух случаях имеет различные профили; Каспаров использует методы, заимствованные им из принципов хорошего дизайна, имевшего место в прошлых играх, поэтому он может распознавать и знать достаточно многое об игре, с тем чтобы проигнорировать огромную часть игрового пространства, через которое Deep Blue вынужден терпеливо продираться *seriatim* (лат.: по зазубринам — *прим. перев.*). На форму поиска, который ведет Каспаров, драматически влияют «озарения», тем не менее это обстоятельство не образует *«совершенно иное»* средство творческого процесса. Всякий раз, когда исчерпывающие поиски Deep Blue блокируют тот *тун* пути,

417

для определения которого у него имеются средства распознавания (трудная, но не невозможная задача), он при возможности может заново использовать R&D, то есть поступить так же, как поступает Каспаров.

Многое из аналитической работы Deep Blue проделано его дизайнерами и придано ему в виде внутреннего багажа. Однако и Каспаров очень многое заимствовал из опыта сотен тысяч часов личного времени, потраченного его предшественниками на исследование шахмат, и знаний, переданных ему игроками в шахматы, тренерами и книгами. В этой связи заслуживает внимание мысль, высказанная недавно Бобби Фишером. Он предлагает воссоздать игру в шахматы в

изначальной чистоте (от информации — *Н. Ю.*) с условием, чтобы в начале каждой игры главные фигуры *случайно* располагались на заднем ряду (случайно, но в зеркальном отражении для черных и белых фигур). Это изменение условий как для людей, так и машин, тотчас же сделает почти совершенно излишними горы запечатленных в памяти шахматных открытий, поскольку в измененной игре только в редких случаях какое-либо из этих профессиональных знаний будет востребовано. Один будет отброшен назад и начнет искать опору в фундаментальных принципах, другому потребуется проделать значительно больше дизайнерской работы в реальное время — до сигнала часов. И далеко не очевидно, даст ли это изменение в правилах игры больше преимуществ человеческим существам или компьютерам. Все зависит от того, какой из этих двух типов игроков в шахматы в конечном итоге в большей мере опирается на механическую память — опору *с минимальным пониманием*, на R&D ранее существовавших искателей.

Фактом является то, что поисковое пространство для шахмат слишком велико для того, чтобы даже Deep Blue мог исчерпывающе его исследовать в реальный отрезок времени. Поэтому, как и Каспаров, он обрезает какие-то поисковые пути и прибегает к риску калькуляции и, как и Каспаров, часто прибегает к этим рискам в довычислительном виде. Можно предположить, что на их очень разных архитектурах оба они

418

прибегают к большому количеству компутаций «грубой силы». В конце концов, что эти самые нейроны знают о шахматах? Любая проделываемая *ими* деятельность должна быть тем или иным видом работы грубой силы.

Кому-то предложенный мною ракурс описания проделанной мозгом Каспарова работы может показаться решением спорного вопроса в пользу вычислительного, то есть предпочтением подхода, характерного для Искусственного Интеллекта. Однако работа должна быть так или иначе проделана, а никакого *другого* способа для понимания проделанной работы никто еще не предложил. Бесплезно говорить, что Каспаров использует «инсайт» или «интуицию», поскольку последнее только означает, что сам Каспаров не имеет никакого привилегированного доступа и никакого понимания того, каким образом он приходит к хорошим результатам. Поскольку никто не знает, как это делает мозг Каспарова, — и меньше всего сам Каспаров, — не существует вообще никакой очевидности в поддержку утверждения, что используемые Каспаровым средства являются «совершенно иными» по сравнению со средствами, которые эксплуатирует Deep Blue. Об этом не нужно забывать, когда возникает соблазн заявлять, что, «разумеется», методы Каспарова совершенно особые. Спрашивается, что же тогда толкает людей к подобному ходу мысли? Принятие желаемого за действительное? Страх?

Вы можете возразить, что это только шахматы, а не искусство. Шахматы *банальны* в сравнении с искусством (это сегодня так считают, когда компьютер играет с чемпионом мира по шахматам). Тогда попробуем обратиться к ЕМІ Дэвида Коупа<sup>1</sup>. Коуп задумал создать простой усилитель эффективности, помогающий композитору снять разного рода помехи при композиции (с ними сталкивается любой творческий человек),

<sup>1</sup> О деталях см.: Cope D. (ed.) *Virtual Music* (MIT Press, forthcoming), куда включен мой комментарий: «Collision Detection, Muselot, and Scribble: Some Reflections on Creativity».

419

или высокотехническую экстензию традиционных поисковых носителей (фортепьяно, нотного стана, магнитофона и т.д.).

По мере роста компетентности ЕМІ, что означает насыщение ее все большим и большим количеством процессов генерирования-и-тестирования, она все больше продвигалась на роль композитора как такового. Когда ЕМІ питают музыкой Баха, она реагирует на это созданием музыкальных композиций в его стиле. Когда ее питают музыкой Моцарта или Шуберта, или Пуччини, или Скотт Джоплин, она с готовностью анализирует их стили, сочиняя новую музыку в соответствии с их духом. Заметим, что ее стилизации оказываются лучше, чем у самого Коупа или почти у любого композитора-человека. Когда ее питают произведениями двух композиторов, она может быстро сочинить музыку, в поразительной манере сведя воедино их стили. А когда ей дают



сразу всю музыку, она (можно сказать, не «отдышавшись») продолжает сочинять, используя всю тотальность полученного ею музыкального опыта.

Результаты композиции, наряду с другой музыкой, создаваемой в формате MIDI, в свою очередь, могут снова и снова вводиться в ЕМІ, и в итоге у нее появляется присущий только ей «персональный» музыкальный стиль, в котором прослеживаются заимствования у мастеров, но в то же время видно, что он, безусловно, является особой интеграцией всего накопленного ею «опыта». Сегодня ЕМІ в состоянии сочинять не только состоящие из двух частей пьесы или песни, но даже целые симфонии и, насколько я знаю, их насчитывается больше тысячи. Они достаточно хороши, чтобы провести глупых экспертов (композиторов и профессоров музыки), и я могу лично засвидетельствовать факт, что прослушивание арии ЕМІ-Пуччини вызвало у меня комок в горле; это была весьма добротная имитация, чтобы обмануть меня. Дэвид Коуп больше не имеет права говорить, что он является композитором симфоний и художественных песнопений ЕМІ. Точно так же, как Мюррей Кэмпбелл не может утверждать, что он победил Каспарова в шахматной игре.

420

В этом новом элементе в каскаде подъемников дарвинист видит только самый последний этап во всей длительной истории. Мы должны признать, что граница между авторами и созданными ими артефактами должна быть столь же проницаема, как все другие границы в этом каскаде. Утверждение Ричарда Докинса, что плотина бобра является такой же частью фенотипа бобра — его *расширенного фенотипа*, — как и зубы, и шкура,<sup>1</sup> создает плацдарм для дальнейшего наблюдения, а именно, что границы человеческого авторства столь же подвержены расширению и продлению. На самом деле все это известно давным-давно, и в течение столетий имелось множество относительно стабильных договоренностей о характере отношения к произведениям творца и его окружения, например, к полотнам Рубенса, к *студии* Рубенса, к работам его учеников. Везде, где появляется рука помощи, мы можем ставить вопрос только о том, кто кому помогает, кто есть творец, а что есть творение.

Как следует относиться к подобным вопросам?

Если антидарвинистам хочется сохранить что-то вроде традиции авторства и иметь какое-то *практическое правило* для определения, кому (или чему) следует воздавать почести (или порицание), которыми сопровождается авторство, их интенции можно принять во внимание (что не обязательно означает, что мы должны исполнять их). Если же им этого недостаточно и они хотят представлять авторов в качестве объективных, имеющих метафизическое основание, представителей «естественного вида» (о, ирония этих волчьих слов эссенциалиста, облаченного в овечью шкуру натуралиста!), значит, они ищут ниспосланное сверху наитие.

Я считаю, что из всего наследия Дарвина отказ от идеи наития свыше для философии является самым глубоким и важным следствием. Его значимость выходит далеко за пределы локальных стычек между эволюционной эпистемологи-

<sup>1</sup> Dawkins R. *The Extended Phenotype*. Oxford and San Francisco, 1982.

421

ей и эволюционной этикой. Принятие «странной инверсии рассуждения» Дарвина требует отказа от завлекательных идей, веками находившихся в центре философской традиции. Не только от эссенциализма Аристотеля и нередуцируемых *целей* (telos), но также от декартовского *res cogitans* — каузальной силы, существующей вне механистического мира. Я указал только на три идеи, казавшиеся особенно несокрушимыми до появления на сцене Дарвина. Песнями сирен этих упорствующих традиций до сих пор заслушиваются философы, не готовые прибегнуть, так сказать, к инверсии. Оставаясь приверженцами до-дарвиновских теорий, они создают для себя проблемы, которые, надо думать, они постараются передать следующим поколениям<sup>1</sup>.

Все эти темы сходятся воедино, когда дело доходит до объяснения творчества и авторства. Здесь у людей возникает потребность схватить «сущность» творчества, докопаться до внутреннего смысла значений и целей, найти местоположение ответственности. При этом предполагается, что все это каким-то образом выходит за рамки каузальной фабрики, в которой они воплощены. Их главная посылка — таланту свойственна способность генерировать осмысленные слова и дела

<sup>1</sup> Приведу три примера: ряд некорректных теорий психосемантики Джерри Фодора; неспособность Джона Сёрля объяснить, каким образом «внутренняя интенциональность» может эволюционировать, когда она не имеет следствий, обладающих «контролирующей силой», видимой для селективного давления; поиск Джоном МакДоуэллом недарвиновской альтернативы тому, что он называет «грубым натурализмом», его борьба за спасение варианта нормативности, являющейся не простой *как если бы* нормативностью, которую он различает в эволюции. Об анализе затруднений, имеющих у Фодора и Сёрля, см.: Dennett D. Darwin's Dangerous Idea, 1995; Dennett D. Granny versus Mother Nature — No Contest // Mind & Language, 11, №3, 1996. P. 263—269; Dennett D. Review of John Searle, «The Rediscovery of the Mind» // Journal of Philosophy, 60 (4), 193—205, Apr. 1993. О моих разногласиях с МакДоуэллом я буду говорить в другой раз.

422

только внутри его *собственных* границ, черпая их из глубин своей собственной *нередуцируемой* гениальности, разительно отличающей нас от простых механизмов<sup>1</sup>.

Платон призывал изображать природу на ее стыках — прекрасный биологический образ, а Дарвин показал, что стыки в биосфере представляют собой не жесткие границы, как это принято считать в теориях эссенциалистов, а наоборот, эмерджентные следствия исторических процессов. По мере того как один вид превращается в два вида, узкие перешейки промежуточных звеньев со временем исчезают, оставляя сгустки-островки, обладающие семейным сходством и окруженные пустотой. Как заметил Дарвин (несколько другими словами), процессы обратной связи усиливают разделение и активно способствуют депопуляции этой срединной почвы. Такой же вид следствий ожидаем и в сфере человеческого сознания и культуры, в сложившихся привычках социума и практики, в которых проявляется тенденция к изоляции процессов создания художественных произведений: они помещаются в одной голове. «Вы являетесь автором этого произведения?» «Все созданное — это ваша собственная работа?»

Простой факт привычности такого рода вопросов свидетельствует о наличии прессы культуры, поощряющего людей *принимать* за истинные нравящиеся им ответы. Маленькая девочка с цветным мелком в руке, торопливо творя свой рисунок, отталкивает протянутую ей руку помощи родителей, брата или сестры, потому что ей хочется, чтобы это был *ее* рисунок. Она уже оценила и приняла норму гордыни авторства — культурно внушенную склонность, запечатленную на территориальном и биологическом праве собственности. Сама идея стать художником формирует ее оценки тех черт, которые она обнаруживает в себе. В свою очередь, эти нормы будут оказывать

<sup>1</sup> Анализ этой темы и поиска Фредом Дретчке привелигированного места, где происходит понимание, см.: Dennett D. Brainchildren. Cambridge MA, 1998. P. 59—80 (раздел: «Do-It-Yourself Understanding»).

423

мощное воздействие на ее способы поиска в Пространстве Дизайна, на ее по большей части подсознательное подражание идеалу Пикассо. Или, если ей присущ дух независимости, она бросит миру вызов, подобно тому, как это сделал Марчел Дучэмп.

Кабанне: Что определяет ваш выбор объектов?

Дучэмп: Это зависит от самого объекта. Вообще-то я стремлюсь остерегаться его «взгляда». Очень трудно выбрать объект, поскольку дней через 15 ты начинаешь либо любить его, либо ненавидеть. Нужно относиться к нему с безразличием, как если бы у тебя отсутствовали эстетические эмоции. Выбор объекта всегда основан на визуальном безразличии и в то же время на полном отсутствии хорошего или плохого вкуса...<sup>1</sup>

Тема управления воображением — одна из постоянных в ведущихся по этому поводу дебатах: стоящие по обе стороны баррикад авторы часто недооценивают ресурсы дарвинизма, не используют все его возможности и выдвигают упрощенные альтернативы. Например, некоторые дарвинисты известны слишком поспешной готовностью найти (или изобрести) различия в *генетической приспособляемости*, которые они затем ищут в каждом найденном ими различии.

В свою очередь, антидарвинисты, указывая на огромную дистанцию между такими разными сотворенными объектами, как улей пчел и «*Страсти Святого Матфея*», склонны утверждать, что всякий, кто попытается объяснить процесс творчества этих двух объектов с помощью одного и

того же набора принципов, повинен в той или иной редуccionистской фантазии. Вроде такой: «У Баха был некий ген для написания контрапункта в стиле барокко, а у пчел наличествует ген для строительства сотов в форме шестиуголь-

<sup>1</sup> *Dialogues with Marcel Duchamp, Pierre Cabanne (transl. Ron Padgett), N.Y., Viking Press, 1971. P. 48.* Моя благодарность Николасу Хэмфри и Виктории МакГир за идеи, выраженные в предыдущем параграфе.

424

ника», или: «Бах был просто бессознательным мутатором метода проб и ошибок и селекции музыкальных мемов, которые к тому времени уже фигурировали в его культурном окружении». Обе альтернативы являются, конечно, чепухой, но фиксация их изъянов никоим образом не служит поддержкой вывода, что («следовательно») при любом объяснении феномена творчества Баха в качестве рабочих следует принимать *не-дарвиновские принципы*, то есть принципы нередуccionизма.

Вместо туманно рисуемой пропасти в виде «Дарвиновского феномена», с одной стороны, и «не-дарвиновского феномена», с другой, нам следует научиться представлять пространство между пчелой и Бахом в виде заполненного разного рода смешанными случаями, едва заметным образом отличающимися от своих ближайших соседей. Научиться заполнять пропасть прослеживаемыми градиентами переходов не-психики (non-minds), прото-психики, геми-деми-семи-психики, психики-сороки, психики-подражателей, психики обезьян, умными психиками, «находящимися в поиске» психиками, «отрывающимися от земли психиками» и, наконец, гениальными разумами. Представлять индивидуальные психики (minds) любого калибра состоящими из разных частей, включая, конечно, имеющие особые цели «модули», адаптированные к различным новым трюкам и задачам, а также каскада приспособлений, необходимых для рефлексий высшего порядка вместе со способностью генерировать все более тонкие и дифференцированные поиски в до-селекционных регионах Огромного пространства возможных дизайнов.

Важно признать, что гениальность как таковая является следствием естественного отбора, включающего в себя процедуры генерирования-и-тестирования, имевшие место на всем пути снизу вверх. Но как только возникает такой результат, больше не имеет смысла рассматривать гениальность исключительно в виде каскада процессов генерирования-и-тестирования. Зачастую есть резоны перескочить на *нарративный* уровень и трактовать агента в виде самости, обладающей

425

разнообразными проектами, целями, предположениями, надеждами... Короче говоря, разумнее принять установку интенциональности по отношению ко всему комплексу плодов эволюционных процессов.

Принятие интенциональной установки позволяет заключить в скобки механические микропроцессы вместе с историей их реализации; они по большей части неизвестны и непознаваемы. Выводя их из фокуса внимания, установка выдвигает на первый план паттерны рациональной деятельности, дорогу которым проложили механические микропроцессы. Тактика интенциональной установки полезна для самопонимания творческой личности. Автору не следует считать удручающим откровением, что при близкой инспекции, даже при близкой *интроспекции*, гениальность растворяется в связке вздора, который, в свою очередь, растворяется в совокупном наборе процессов проб-и-ошибок, над которым никто не властен.

Означает ли, что такое понимание равносильно потере — или элиминации — самости, гениальности, творчества? Хорошо знакомые с процессом творчества и рефлексиирующие по этому поводу люди творческих профессий — великие художники и научные гении — в общем-то спокойно относятся к такому открытию. Говорят, что Моцарт по поводу рождения его лучших музыкальных идей заметил: «Откуда и как они приходят? Я не знаю и не имею никакого отношения к этому»<sup>1</sup>. Художник Филипп Гастон говорит, что, когда его обуревают жажда творчества, его не волнует испарение видимой самости.

«Когда я вхожу работать в студию, в ней роится шумная и досаждающая мне толпа; мелькают все

известные в истории живописцы, все мои современники, все художественные критики и т.д. По мере того как я включаюсь в работу, они один за другим покидают меня. Если день оказался удачным, они

<sup>1</sup> В часто цитируемом, хотя, возможно, и не подлинном пассаже см.: D. Dennett. *Darwin's Dangerous Idea*. P. 346—347.

426

исчезнут все до одного. Если меня действительно посетила удача, я тоже исчезаю<sup>1</sup>.

В заключение я хотел бы высказать признательность нескольким моим соавторам. Анонимно.

Хорхе Луис Борхес

Дэвид Коуп

Чарльз Дарвин

Ричард Докинс

Сьюзен Деннет

Рене Декарт

Марсель Дучэмп

Томас Эдисон

Бобби Фишер

Филип Гастон

Дуглас Хофстадтер

Николас Хэмфри

Роберт МакКензи

Тони Марсель

Виктория Макгир

Людвиг Мис ван дер Роэ

Пабло Пикассо

Уильям Паундстон

Джон Сёрль

Уильям Шекспир

Мэри Шелли

Поль Валери.

<sup>1</sup> Я не смог установить источник цитаты Гастона, но я нашел очень похожее замечание, которое приписывают композитору Джону Кэйджу, современнику и близкому другу Гастона, который [как говорят] высказался о его рисунках так: «Когда ты работаешь, все на свете находится в твоей студии — прошлое, твои друзья, мир искусства и, помимо всего прочего, твои собственные идеи. Но если ты продолжишь писать, все это начнет уходить от тебя одно за другим, и ты останешься совсем один. И если тебе повезет, сам ты тоже уйдешь». (Гастон и я, как и все другие творческие люди, тоже любим заново использовать то, что мы обнаруживаем, делая время от времени небольшие добавления.)

427

## 2. Почему каждый из нас является новеллистом

Daniel C. Dennett. *Why Everyone is a Novelist* // *Times Literary Supplement*. September 16—22, 1988. P. 96—98.

Что такое самость? На этот вопрос я постараюсь ответить, прибегнув к аналогии с чем-то более простым и повсеместным, столь же загадочным, как и самость, и обладающим некоторыми общими с ней свойствами.

Я имею в виду *центр гравитации* объекта. Данное понятие прекрасно работает в физике Ньютона. Центр гравитации не является ни атомом, ни субатомной частицей, ни какой-либо иной физической единицей мира. У него нет массы, нет цвета и вообще никаких физических свойств, кроме пространственно-временной локализации. В этом отношении он может служить прекрасным примером того, что Ганс Рейхенбах назвал бы *abstractum*. Это чисто абстрактный

объект. Если хотите — изобретенная теоретиком фикция (fiction). Центр гравитации нельзя считать реальной, дополнительной к атомам вещью во Вселенной. Тем не менее в рамках физики роль этой фикции тщательно определена, подробно описана и хорошо работает.

Позвольте мне напомнить, сколь упрямой и знакомой является идея центра гравитации. Возьмем, к примеру, стул. Как и у прочих физических объектов, у него имеется свой центр гравитации. Начните потихоньку наклонять его и вы с большей или меньшей точностью сможете предсказать его падение или, по прекращении вашего воздействия на него, возвращение в прежнее положение. Все мы являемся хорошими предсказателями поведения центра гравитации и изобретателями объяснений, когда и почему вещи падают. Положите книгу на стул. У нее тоже есть центр тяжести. Двигая ее к краю, вы знаете, что в какой-то точке она упадет. Упадет тогда,

428

когда ее центр тяжести перестанет точно соответствовать точке ее поддерживающей основы (в сиденье стула). Заметьте, что данное суждение само по себе по сути является тавтологическим; все ключевые термины в нем взаимоопределяемы. Тем не менее оно также может фигурировать в объяснениях, выглядящих как некоторая разновидность каузальных объяснений. На вопрос «Почему эта лампа не опрокидывается?» следует ответ: «Потому, что ее центр тяжести находится достаточно низко». Относится ли данный ответ к каузальному объяснению? Во всяком случае, он может конкурировать с явными каузальными объяснениями типа: «Потому, что она пригнана гвоздем к столу» или «Потому, что она поддерживается проводами».

Центрами тяжести можно манипулировать. Например, я с легкостью изменю центр тяжести кувшина, отлив из него какую-то часть воды. Поэтому, хотя центр тяжести есть чисто абстрактный объект, ему присуще пространственно-временное перемещение, на которое я могу воздействовать. У него есть история, однако она может включать некоторые весьма странные эпизоды. Перемещаясь в пространстве и времени, движение центра тяжести может быть прерывным. Взяв кусочек жевательной резинки и быстро приклеив его к ручке кувшина, мы своим действием сможем сдвинуть центр тяжести кувшина с точки А в точку Б. При этом ему нет необходимости перемещаться через все промежуточные положения. Будучи abstractum, он не скован всеми ограничениями физического путешествия.

Обратимся теперь к центру тяжести немного более сложного объекта. Предположим, что нам нужно проследить путь перемещения центра тяжести в какой-нибудь сложной машине, имеющей множество переключений скоростей, распределительных валов, рычагов обратной связи и т.д., — скажем, в машине с паровым двигателем. Предположим далее, что наша теория относительно операции с машиной позволяет нам точно проследить в ней сложную траекторию центра тяжести. Предположим также совершенно невероятное: мы обнаружим,

429

*что в этой именно* машине траектория центра тяжести является точно такой же, как и траектория отдельного атома в коленчатом валу. Тем не менее, даже обнаружив совпадение, мы ввели бы себя в глубокое заблуждение, придерживаясь гипотезы о совпадении (тождественности) центра гравитации этой машины с данным атомом железа. К тому же допустили бы серьезную категориальную ошибку, ибо центр тяжести *не что иное*, как abstractum. Он только воображаемый объект. Относя его к категории воображаемых объектов, я не имею намерения тем самым принизить его; он — замечательный воображаемый объект и совершенно законно занимает место внутри серьезной, трезвой, echt (подлинной — *Н. Ю.*) науки физики.

Самость также представляет собой абстрактный объект и тоже порождение теоретика. Теория самости не есть теория частиц физики; скорее, ее следует отнести к тому, что можно было бы назвать разновидностью народной физики (people-physics)<sup>1</sup>. Она известна в виде феноменологии, или герменевтики, или науки о душе (Geisteswissenschaft). Физик занимается интерпретацией, если хотите, стула и его поведения и изобретает теоретическую абстракцию центра гравитации, которая затем становится весьма полезной при определении будущего поведения стула в самых разнообразных условиях. Герменевтик или феноменолог — или антрополог — в качестве своего объекта выбирает некоторые значительно более сложные вещи, движущиеся в мире (бытия



человеческие и животных), но сталкивается со сходной проблемой интерпретации. Ему представляется, что теоретическую ясность можно получить при организации интерпретации вокруг центральной абстракции: каждая личность наделена самостью (в виде добавки к центру гравитации). На деле же в этом случае мы должны также полагать самости за основание наших собственных самостей. Заметим, что теоретическая проблема само-

<sup>1</sup> Термином «reople-physics» обозначают обыденные, в отличие от научных, представления о физическом мире.

430

интерпретации является не менее трудной и важной, нежели проблема интерпретаций других. А теперь зададимся вопросом, насколько самость отличается от центра гравитации? Прежде всего отметим, что самость — значительно более сложное понятие. Я попытаюсь прояснить его, прибегнув к аналогии с другой разновидностью воображаемых объектов — вымышленными (fictional) персонажами литературы. Возьмите повесть «Моби Дик» Германа Мелвилла и откройте ее на первой странице. Там говорится: «Называй меня Ишмаелем». Кого называть Ишмаелем? Называть Мелвилла Ишмаелем? Нет. Называйте Ишмаеля Ишмаелем. Писатель Мелвилл придумал воображаемого персонажа и назвал его Ишмаелем. При чтении этой книги вы получите сведения об Ишмаеле, его верованиях и желаниях, поступках и пристрастиях. Однако из повести вы извлекаете значительно больше об Ишмаеле, нежели Мелвилл сообщает вам открытым текстом. Что-то из этого большего вы в состоянии вычитать из текста путем импликации, что-то — путем экстраполяции. Однако за границами такой экстраполяции воображаемые миры остаются просто неопределенными. К примеру, зададимся таким вопросом: «Было ли у Шерлока Холмса три ноздри?»<sup>1</sup>. Ответ, конечно, будет «нет» не в силу того, что Конан-Дойль где-нибудь высказался по этому поводу или специально пояснил наличие у Шерлока Холмса двух ноздрей, а потому, что мы вправе делать такую экстраполяцию. При отсутствии свидетельств о противоположном, вполне резонно предположение о нормальном носе у Шерлока Холмса. Зададимся другим вопросом: «Была ли на левой лопатке Шерлока Холмса родинка?». Ответ будет и «да», и «нет». Потому что ничто в тексте или в принципах экстраполяции текста не позволяет нам ответить на этот вопрос. Просто нет ни одного факта, относящегося к нему. Почему? Потому что Шерлок Холмс не что иное, как воображаемый персонаж, со-

<sup>1</sup> Пример Дэвида Льюиса (David Lewis. Truth and Fiction // American Philosophical Quarterly, 1978).

431

зданный или конституированный текстом и культурой, в которой этот текст фигурирует. Неопределенность является фундаментальным свойством воображаемых объектов, и этим они резко отличаются от вида объектов, которыми оперируют ученые: теоретических сущностей или того, что Рейхенбах называл *illata* — от *возможно* выводимых сущностей, таких как атомы, молекулы и нейтрино. По этому поводу логик мог бы сказать, что «принцип двузначности»<sup>1</sup> не распространяется на воображаемые объекты. Иначе говоря, по отношению к любому действительному человеку, живому или мертвому, вопрос о наличии или отсутствии у него родинки на левой лопатке имеет ответ — либо «да», либо «нет». Была ли у Аристотеля такая родинка? Даже если нам никогда не удастся обнаружить действительный факт, он существует. В то время как по отношению к воображаемому персонажу этот вопрос может вообще не иметь ответа.

Далее представим, скажем, некоего невежественного литературного критика, не понимающего, что воображаемое есть только воображаемое. Он придерживается странной теории о работе воображаемого; он думает, что в процессе написания писателем повести происходит что-то буквально [поистине] магическое. Когда тот наносит слова на бумагу, критик говорит (такие утверждения часто можно услышать, хотя они не воспринимаются слишком буквально), что писатель поистине творит мир. Лакмусовой бумажкой в тестировании столь эксцентричного взгляда является принцип двузначности. Когда наш предполагаемый критик говорит о воображаемом мире, он имеет в виду странный тип реального мира, на который распространяется принцип двузначности. Такой критик мог бы всерьез удивляться по поводу того, был ли доктор

Ватсон вторым кузеном Мориарти, или был ли кондуктор поезда, сопровождавший Холмса и Ватсона в Алдершот, тем же самым кондуктором поезда, с каким они возвращались назад в Лон-

<sup>1</sup> В логике принцип двузначности означает либо «да», либо «нет», т.е. либо факт есть, либо его нет.

432

дон. Конечно, такого рода вопросы не могут возникать при правильном понимании воображаемого. В то же время аналогичные вопросы по отношению к реальным историческим персонажам должны иметь ответом «да» или «нет», даже если нам никогда не удастся извлечь их из прошлого.

Центры гравитации, будучи фиктивными объектами, демонстрируют такую же особенность. У них имеются только те свойства, которыми их наделяет конституирующая теория. Если, почесавшись в затылке, вы скажете: «А может быть, центры гравитации в действительности являются нейтрино!», — вы окажетесь жертвой заблуждения относительно теоретического статуса центра гравитации.

Теперь обратимся к нашему главному вопросу: как я могу утверждать, что самость — ваша реальная самость — является чем-то вроде воображаемого персонажа? Разве не зависит творение всех воображаемых самостей от существования реальных самостей? Может показаться, что дело обстоит именно так, но я постараюсь убедить, что такое представление иллюзорно. Вернемся к Ишмаелю. Ишмаель — воображаемый персонаж, хотя мы можем, конечно, все узнать о нем. Кто-то посчитает его даже более реальным, нежели многих из своих друзей. При этом его рассуждения будут строиться следующим образом: Ишмаель был придуман писателем Мелвиллом, а Мелвилл — реальный персонаж, то есть был реальным персонажем. Реальной самостью. Разве это не свидетельство того, что для сотворения воображаемой самости требуется существование реальной самости? С моей точки зрения, не требуется, однако, чтобы убедить вас в этом, мне нужно провести ваше воображение через одно упражнение.

Для начала предложу вам представить нечто, что некоторые из вас посчитают за невероятное, — машину, пишущую художественные повести. Вы можете относиться к ней как к продукту исследований по искусственному интеллекту или компьютеру, созданному или запрограммированному для написания повестей. Однако проект машины не ограничен написанием какой-либо определенной повести. Мы можем предположить

433

(если это поможет), что ей придан огромный массив необходимой информации, которая ей, возможно, потребуется для написания повести, а также некоторые частично случайные и отсюда непредсказуемые способы завязки и построения сюжета. Представим теперь, что вокруг машины собрались ее проектировщики и любопытствуют, какого рода повесть намерена написать созданная ими машина. Они включают ее, через какое-то время высокоскоростной принтер защелкает и выдает первую фразу: «Называй меня Гилбертом». Далее последует явно автобиография некоего воображаемого Гилберта. Итак, Гилберт — это воображаемая, сотворенная самость, однако она сотворена не какой-либо самостью. Конечно, существовали специалисты — вполне реальные люди, спроектировавшие машину, тем не менее не они сотворили Гилберта. Гилберт — продукт процесса дизайна или изобретения, в котором вообще нет какой-либо самости. Я намеренно оговариваю, что данная машина не является мыслящей, она — не «мыслитель». Она представляет собой немую машину, наделенную способностью писать сносные повести. (Если вы находите такую идею совершенно невероятной, я попрошу вас представить доказательства в пользу вашей точки зрения и посоветую продолжить чтение моей статьи; закончив, вы, возможно, потеряете интерес к защите вашего рискованного заявления о невозможности написания повести машиной.)

Итак, представим себе, что нам удалось извлечь из машины некую сносную повесть. Заметим, что данная повесть может быть подвергнута такому же типу толкования, какой мы обычно применяем по отношению к любым другим литературным произведениям. Если вам понадобилось взять в библиотеке повесть — по случайному выбору, на практике вы не смогли бы сказать с

уверенностью, что она не написана чем-то вроде этой машины. (А если вы — Новый Критик, вас это и не будет волновать.) Взяв текст, вы сможете его интерпретировать, а следовательно, узнать об истории, жизни и приключениях Гилберта. По мере вашего

434

чтения и ваши ожидания, и предсказания, и ваша интерпретативная реконструкция уже прочитанного будут прикованы к центральному узлу повести — воображаемому персонажу Гилберту.

Теперь я хочу несколько сместить ход этого мыслительного эксперимента. Итак, мы вообразили повесть «Жизнь и времена Гилберта», которую постукивая выдал компьютер, являющийся ни чем иным, как ящиком, стоящим в углу какой-нибудь лаборатории. Далее я предлагаю внести еще одно дополнение в наш мысленный эксперимент с компьютером и представить, что у него имеются ноги и руки, или лучше — колеса. (Мне не хочется делать его слишком антропоморфным.) Его снабдили телевизионным глазом и способностью двигаться по миру. Свое повествование он также начинает словами «Зови меня Гилбертом» и рассказывает свою повесть. Однако сейчас мы замечаем, что если прибегнем к уловке, к которой, по мнению Новых Критиков, мы никогда не должны прибегать, и посмотрим за пространство текста, то обнаружим, что в реальном мире существует истинно-сохраняющая интерпретация этого текста. Приключения воображаемого персонажа Гилберта имеют удивительное отношение, не являющееся, скорее всего, случайным стечением обстоятельств, к приключениям данного робота, описывающего круги вокруг мира. Если вы ударите по роботу бейсбольной битой, вскоре после этого в историю Гилберта будет включен эпизод, что кто-то, выглядящий именно так, как вы, ударил его бейсбольной битой. Робот запрется в чулане и затем молвит: «Помоги мне!» Помочь кому? Ну, хорошо, можно предположить, что Гилберту. Но кто такой Гилберт? Является ли Гилберт роботом или просто созданной роботом воображаемой самостью? Если мы подойдем и поможем роботу выйти из чулана, он пошлет нам записку: «Благодарю. Люблю. Гилберт».

В данном случае мы уже не можем игнорировать факт, что в воображаемой карьере воображаемого Гилберта имеется интересное сходство с «карьерой» нашего простого двигающегося по миру робота. Мы все еще можем утверждать, что

435

мозг робота — компьютер робота — в действительности ничего не знает о мире; это не самость. Перед нами всего-навсего шелкающий компьютер. Ему неизвестно, что он творит. Он не знает даже о своем собственном творении воображаемого персонажа. (Заметим в связи с этим, что то же самое можно сказать и о вашем мозге; ему вообще неизвестно, что он делает.) Тем не менее контролируемый компьютером поведенческий паттерн в принципе может быть интерпретирован нами как разрастающаяся биография или нарративное повествование о самости. При этом мы — не единственные интерпретаторы. Робот-рассказчик тоже, конечно, является интерпретатором; самоинтерпретатором, выдающим собственное объяснение своей деятельности в мире.

К данной аналогии, как мне представляется, следует отнестись со всей серьезностью. У философа-материалиста или ученого, занимающегося высшей нервной деятельностью человека, может возникнуть вопрос: «Где же находится самость?». Начинать искать ответ на него с поиска самости в мозгу — значит допускать категориальную ошибку. В отличие от физических центров гравитации, чьим единственным свойством является их пространственно-временное положение, у самостей их пространственно-временное положение определено только в весьма общих чертах. Проще говоря, в обычной ситуации, если на скамейке парка сидят три человеческих существа, значит, имеются три самости, расположенные в ряд и приблизительно на одном расстоянии от созерцаемого ими фонтана. Или, прибегнув к несколько устаревшей фразеологии, можно говорить о количестве душ, находящихся в парке. (Вспомните обычное клише: «Все двадцать душ, находившиеся на правом борту спасательного судна, были спасены, а оставшиеся на палубе пропали».)

Исследования мозга смогут позволить нам сделать некоторые более точные локализации, однако возможность получения их еще не дает основания для предположения, что процесс локализации

будет продолжаться бесконечно и что однажды придет день, когда мы сможем сказать: «Эта клеточ-

436

ка — в центре hippocampus (или где-нибудь еще) — как раз и есть искомая самость!».

Конечно, существуют большие различия между воображаемыми персонажами и нашими собственными самостями. Одно из них, на которое я хотел бы обратить внимание, состоит в том, что воображаемый персонаж обычно предстает как *fait accompli* (фр.: нечто свершившееся — *Н. Ю.*). Вы читаете повесть после ее написания и публикации. На этом этапе писателю-новеллисту слишком поздно сделать определенным что-то неопределенное в сюжете, вызвавшее у вас недоумение или заинтересовавшее вас. Достоевский мертв, и вы не можете узнать у него, о чем еще думал Раскольников, сидя в полицейском участке. Однако повести и не должны быть такими. Джон Апдайк написал три повести о Кролике Энгстроме: «*Кролик бежит*», «*Кролик возвращается*» и «*Кролик разбогател*». Предположим, что читатели, которым особенно понравилась первая повесть, соберутся вместе и составят список вопросов Апдайку — относительно разных вещей, которые, как бы им того хотелось, ему следовало описать в первой повести, когда Кролик был молодым человеком, уже побывавшим баскетбольной звездой. Они могли бы послать свои вопросы Апдайку с просьбой подумать о другом рассказе в этой серии, но только не придерживаться в этом случае хронологического следования. Как и в случае с «Александрийским квартетом» Лоуренса Дюрреля, серия рассказов о Кролике могла бы включить новеллу о ранних годах Кролика, когда он еще играл в баскетбол, и в ней содержались бы ответы на вопросы читателей.

Замечу только, какие действия в этом случае не следует предпринимать. Не следует говорить Апдайку: «Ответьте нам на вопросы, которые вам уже известны, и сообщите уже имеющиеся у вас сведения. Продолжите вашу работу и позвольте нам узнать о всех утаенных от нас секретах». Не следует также просить Апдайка предпринять исследование, уместное в отношении автора многотомной биографии реальной исторической личности. Апдайку можно было бы попросить

437

написать новую новеллу, изобрести по нашему желанию еще несколько новых новелл. В случае, если бы он согласился пойти нам навстречу, в процессе написания нового рассказа он, возможно, расширил бы и сделал более определенным персонаж Кролика Энгстрома. Благодаря его творческим усилиям, до этого не проясненные обстоятельства, возможно, станут более определенными.

Выскажу предположение, что данное воображаемое упражнение с Апдайком, которого мы заставляем написать еще какие-то рассказы для ответа на наши вопросы, на самом деле нам хорошо знакомо. Именно таким образом мы рассматриваем друг друга; это способ узнать, какие мы есть. Мы не в состоянии переделать уже завершившиеся и определившиеся куски нашего прошлого, но мы постоянно делаем все более определенными наши самости по мере реагирования на способы, какими мир наносит нам удары (*impinges*). Конечно, личность способна вовлекаться в аутогерменевтику и интерпретацию своей собственной самости, в частности, может оглянуться назад и подумать о своем прошлом, о своих воспоминаниях, заново передумать их и переписать. Данный процесс изменяет «воображаемый» персонаж — тот характер, каким вы являетесь, — во многом таким же манером, каким Кролик Энгстром, после того как Апдайк написал бы вторую новеллу о нем как о молодом человеке и он оказался бы существенно иным воображаемым персонажем, наделенным ранее отсутствовавшими чертами. Все это было бы в высшей степени мистическим и магическим проектом (и отсюда чем-то, что никому не следует принимать всерьез), если бы самость не была бы ничем иным, как *abstractum*.

Для пояснения хода моего рассуждения выделю еще один момент в мысленном эксперименте с Апдайком. Апдайк мог бы пойти навстречу нашим пожеланиям и затем забыть их. В конце концов после написания повести «Кролик бежит» прошло немало лет. Возможно, у него не возникло желание возвращаться к старой работе и заново тщательно ее перечитывать; но и в том случае, если бы он вдруг написал новую

438

новеллу, вполне может оказаться, что она противоречила бы первой. Например, Кролик оказался одновременно в двух местах. Если же мы вздумаем выяснить, какая из этих историй была истинной, мы допустим ошибку, ибо истинной истории не существует. При таких условиях выяснять истину было бы просто нарушением когерентности всех имеющихся в вашем распоряжении данных о Кролике. Поскольку Кролик является воображаемым персонажем, нам не следует бить себя по лбу и с удивлением восклицать: «Боже мой! Оказывается, во Вселенной существует раскол; мы обнаружили противоречие в природе!» Имея дело с художественным вымыслом, нет ничего проще противоречия; воображаемый персонаж может иметь противоречивые свойства именно потому, что он воображаемый. Тем не менее мы считаем такие противоречия нетерпимыми при попытке интерпретации чего-либо или кого-либо, даже воображаемых персонажей, поэтому для разрешения конфликта мы обычно прибегаем к раздвоению их характера.

В редких случаях и с реальными людьми происходит что-то вроде такого раздвоения. Рассмотрим предположительно правдивые истории, собранные в рассказе К. Тигпена и Х. Клекли «*Три лица Евы*» (1957) и в рассказе Ф.Р. Шрейбера «*Сибил*»<sup>1</sup>. Три лица Евы, по-видимому, были лицами трех различных личностей; а женщина, описываемая в рассказе «*Sybil*», по всей видимости, имела множество различных самостей, во всяком случае так все это выглядит в рассказах. Как можно осмысленно представить всю эту ситуацию? Для этого имеется только один путь — трезвый скептический способ, применяемый некоторыми психотерапевтами, с которыми я беседовал о такого рода аномалиях. Когда Сибил впервые посетила своего психотерапевта, она не была несколькими разными людьми, запрятанными в одном теле; она была машиной, пишущей новеллы и попавшейся на удочку въедливо-

<sup>1</sup> Thigpen C.H. and Cleckly H. *The Three Faces of Eve* (1957); Schreiber F.R. *Sybil* (1973).

439

го читателя-терапевта, задающего весьма хитроумные вопросы. Вместе с терапевтом они составили невинный союз для написания множества глав новой новеллы. И конечно, поскольку Сибил была чем-то вроде живой новеллы, она вышла в мир и вовлеклась в него с этими новыми самостями, созданными более или менее по требованию и при ревностных советах терапевта.

Сейчас я думаю, что данная идея чрезмерно скептична. Популяционный взрыв новых персонажей, обычно сопровождающий начало психотерапии для страдающих от расстройства, известного как «множественная личность» (Multiple Personality Disorder), вероятно, следует объяснять именно таким образом; однако в ряде случаев есть вполне убедительные свидетельства того, что некоторое умножение самостей (скажем, две, три или четыре) закладывалось уже в начале биографии и до того, как появился терапевт и стал что-то «вычитывать». Во всяком случае Сибил является удивительно патологическим примером чего-то вполне нормального, того паттерна поведения, который мы можем найти в себе. Все мы являемся болтливими существами, рассказывающими и заново пересказывающими себе историю нашей собственной жизни, не обращая особого внимания на вопрос об ее истинности. Почему же мы ведем себя таким образом? Почему все мы являемся столь закоренелыми и изобретательными автобиографическими новеллистами? Умберто Матурана как-то заметил (непротиворечиво): «Все сказанное сказано одним говорящим другому говорящему, которым можете оказаться вы сами». Но почему кто-то должен говорить с самим собой? Почему болтовня не является в высшей степени праздным делом, столь же систематически бесполезным, как и попытки приподнять себя за собственные шнурки?

Главный ключ к объяснению можно найти в том виде феномена, который был открыт Майклом Гэзанига<sup>1</sup> при ис-

<sup>1</sup> М. Гэзанига (Michael Gazzaniga) — специалист в области когнитивных наук {прим. перев.}.

440

следовании редко встречающихся индивидов — «субъектов с рассеченными мозгами», чей corpus callosum был разъединен хирургически, что создало у них две независимые кортикальные полусферы, которые в каких-то *случаях* могут получать разную информацию о происходящем.



Произошел ли в результате операции раскол одной самости на две? Обычно после операции пациенты не проявляют признаков психологического раскола и выглядят не меньшим единством, нежели вы или я, за исключением случаев, когда для них создаются некоторые специально придуманные и хитроумные условия. С точки зрения Гэзанига, поведение пациентов свидетельствует не столько о сохранении единства, существовавшего у них до хирургического вмешательства, сколько о том, что единство нормальной жизни есть иллюзия.

В нормальном сознании Гэзанига видит не красиво упакованное единство, а скорее проблематично сопряженную между собой связку частично автономных систем. Системы сознания не являются в любое время и в равной мере доступными друг другу. У этих модулей или систем временами возникают проблемы с внутренними коммуникациями, которые они разрешают различными изобретательными и окольными путями. Если его объяснение верно (а я думаю, что это так), оно может подсказать ответ на самый загадочный вопрос об осознанном мышлении: что же оно такое? Данный вопрос требует эволюционного ответа и он, конечно, должен быть спекулятивным. (Для моего спекулятивного ответа в настоящее время не является критическим вопрос о том, где внезапно прерывается генетическая эволюция и трансмиссия и ее эстафету принимает культурная эволюция и трансмиссия.)

Согласно Джулиану Джейнсу<sup>1</sup>, объяснение которого я приемлю, сначала появились говорившие — наши предки, еще не являвшиеся по-настоящему осознающими существами. Они говорили, но являли собой именно вид болтливых существ,

<sup>1</sup> Jaynes J. *The Origins of Consciousness in the Breakdown of the Bicameral Mind*. 1976.

441

чья болтливость более или менее схожа с танцами пчел или с разговором компьютеров друг с другом. Конечно, такого рода разговоры не были еще осознанной коммуникацией. Иногда, сталкиваясь с какими-то проблемами, наши предки «звывали» о помощи (что более или менее схоже с ситуацией, когда закрытый в чулан Гилберт молвит: «Помоги мне!»). Временами, возможно, рядом оказывался кто-то, слышащий призыв. Так у прародителей сложилась привычка просить о помощи и, в частности, задавать вопросы. Всякий раз, не находя разрешения какой-то проблемной ситуации, они задавали вопрос, не адресованный к какому-то определенному существу, и временами кто-то, стоявший рядом, отвечал на него. Их дизайн оказался приспособленным, для того чтобы, когда их спрашивали, быть провоцируемым отвечать на такого рода вопросы, — что явилось наивысшим проявлением их способностей.

И вот однажды один из наших далеких предков, находясь в неблагоприятной ситуации, задал вопрос тогда, когда вокруг него не было слушателей. Удивительно, но он услышал свой собственный вопрос, и это стимулировало его подумать о нем совместно с ответом, и, конечно, ответ пришел к нему. Не осознавая проделанного, он установил коммуникационную связь между двумя частями своего мозга, доступ которых друг к другу по ряду глубоких биологических причин был проблематичен. Одна составляющая сознания сталкивается с проблемой, которую в состоянии решать другая составляющая, если только проблема могла быть поставлена перед последней. Благодаря сложившейся привычке задавать вопросы, наш предок наткнулся на маршрут, идущий через уши. Что за открытие! Иногда разговор с самим собой и выслушивание самого себя дает поразительные эффекты, не достижимые иными способами. Для придания смысла данной идее требуется только гипотеза, что модули сознания имеют различные возможности и способы осуществления своей работы и в этом отношении не являются в полной мере взаимодоступными. В этих условиях может оказаться истинным предположение, что способ заставить вас заниматься проблемой — это мучить ею ваше

442

ухо, побуждая искать ее решение ту часть вашего сознания, для которой лучшим стимулом является постановка вопросов. И порою вы обнаруживаете искомый ответ на кончике вашего языка.

Этого было достаточно, чтобы установить эволюционное подкрепление поведению (которое также может быть только культурно передаваемым), проявляющемуся в разговоре с самим собой. Как

отмечали многие авторы, осознающее себя мышление выглядит — во всяком случае большая часть его — в виде разнообразия особенно эффективных и частных разговоров с самим собой. Эволюционный переход к мысли тогда становится более объяснимым. Нам следует только предположить, что путь, цикл, который сначала проходит через рот и ухо, становится короче. Люди «пришли к осознанию», что на практике вокализация и аудиция — весьма неэффективная часть петли. Кроме того, если вокруг вас находились люди, которые могли подслушать произносимое вами, вы могли сбросить больше информации, нежели вам того хотелось. В результате всего этого была развита привычка к субвокализации, что, в свою очередь, могло стать главным руслом к появлению сознательной вербальной мысли.

Джилберт Райл в посмертно изданной книге «*О мышлении*» (Ryle J. On Thinking, 1979) задал вопрос: «Что делает Le Penseur?» Для бихевиориста, каким является Райл, это реальная проблема. Одна часть подбородка-на-кулаке-с-напряженным-лбом выглядит точно такой же, как и другая часть, тем не менее какая-то часть приходит к правильным ответам, а какая-то нет. Что может происходить в этом случае? Парадоксально, что у архибихевиориста Райла есть некоторые весьма слабые наметки объяснения. Осознающую себя мысль, утверждал Райл, следует понимать по модели самообучения, а еще лучше — само-школы или наработки навыков. О том, как на деле должно происходить такое самообучение, Райл говорит весьма скупно, хотя из сказанного им можно получить некоторое первоначальное понимание, а именно, что мы не являемся капитанами наших кораблей; не существует осознающей

443

самости, бесппроблемно управляющей ресурсами сознания. Скорее всего, мы представляем собой в какой-то мере разъединенное единство. Составляющие нас модули должны действовать оппортунистически и вместе с тем поразительно плодотворными способами, чтобы произвести модикум поведенческого единства, которое затем усиливается иллюзией большого единства.

Проведенное М. Гэзанига исследование показывает, иногда с живыми деталями, как именно это должно происходить. Рассмотрим некоторые из его примеров, свидетельствующих об экстраординарной изобретательности правого полушария (чего-то в нем) при столкновении с проблемой коммуникации. В одной серии экспериментов субъектам, у которых до этого был рассечен мозг, предлагалось протянуть левую руку в закрытую сумку, ощупать находящийся там объект и вербально его идентифицировать. Известно, что тактильные чувствительные нервы левой руки ведут в правое полушарие, в то время как контроль за речью обычно находится в левом полушарии, и для большинства из нас вербализация не составляет проблемы. У нормальной личности левая рука знает, что делает правая рука, благодаря *corpus collosum* мозга, поддерживающего систему взаимной информированности обоих полушарий. У субъектов с рассеченным мозгом эта объединяющая связь разрушена; правое полушарие получает информацию об ощущаемом объекте от левой руки, а левому полушарию, осуществляющему контроль за языком, надлежит сделать идентификацию публичной. Однако получается так, что «часть, которая говорит», находится во мраке, в то время как «часть, которая знает», не может сделать ее знание публичным.

Тем не менее находится кружной путь преодоления этой трудности и, согласно наблюдениям, пациентам с рассеченными мозгами удается найти его. В то время как обычные тактильные ощущения представлены противоположно (*contralaterally*) — сигналы идут в противоположное полушарие, — болевые сигналы представлены также и окольным путем (*ipsilaterally*). То есть, благодаря характеру разветвления нерв-

444

ной системы, болевые сигналы идут к обоим полушариям. Предположим, что объект, находящийся в сумке, — карандаш. Правое полушарие временами будет прибегать к очень умной тактике: подсказывать держать карандаш в левой руке таким образом, чтобы его острый грифель с силой надавливал на подушечку пальцев; вызванная нажимом боль даст знать левому полушарию, что в сумке находится что-то острое, и этого окажется достаточным, чтобы начать гадать; правое полушарие будет сигнализировать «становится теплее» и «схватит остроту» улыбкой или другими

контролируемыми знаками, и через весьма короткое время «субъект» — кажущийся «единый обитатель» тела — будет в состоянии вербально произнести правильный ответ.

У нас имеются следующие гипотезы: либо у субъектов с рассеченным мозгом этот экстраординарный кружной талант развился вследствие операции, столкнувшей их с радикальной проблемой доступа полушарий друг к другу, либо операция выявила, но не создала виртуозный талант, свойственный также и нормальным людям. И конечно, Гэзанига утверждает, что для исследования последняя гипотеза предпочтительнее. Иначе говоря, по-видимому, все мы являемся виртуозными новеллистами, вовлеченными во все более или менее унифицированные виды поведения, но временами и в дизунифицированные, и мы всегда, если удастся, делаем вид, что все в порядке. Мы пытаемся превратить все наши материальные сцепления в единую хорошую историю. И эта история — наша автобиография.

Главный воображаемый персонаж в центре этой автобиографии — чья-то самость. Если вам все еще хочется знать, чем же реально является эта самость, вы совершаете категориальную ошибку. В конце концов, когда система поведенческого контроля человеческого существа серьезно повреждена, может оказаться, что наилучшая герменевтическая история, которую мы можем рассказать об этом индивиде, сведется к тому, что в этом теле «обитает» более чем один персонаж. Такое вполне возможно, если исходить из предложенного мною

445

здесь взгляда на самость; исходя из него, не требуется прибегать к каким-либо фантастическим метафизическим тайнам. Можно обнаружить, что наличие множественных самостей в личности столь же непроблематично, как непроблематично найти Раннего Молодого Кролика и Позднего Молодого Кролика в воображаемых новеллах Апдайка: в этом случае следует только допустить, что история не завязана вокруг единой самости, одной воображаемой точки, но завязана (во всяком случае, лучше завязана) вокруг двух различных воображаемых точек.

Временами нам приходится сталкиваться с психологическими неполадками у людей или разъединением связей в результате хирургического вмешательства, когда единственный способ их интерпретации или осмысления состоит в допущении существования двух центров гравитации, двух самостей. Поступая таким образом, мы не создаем и не открываем маленький кусочек духовного вещества. Мы просто создаем другую абстракцию. Она используется как часть теоретического аппарата для понимания, предсказания и осмысления поведения некоторых очень сложных вещей. Тот факт, что эти абстрактные самости кажутся нам столь настойчиво присутствующими и реальными, неудивителен. Они являются значительно более сложными теоретическими сущностями, нежели центры гравитации. Напомню в связи с этим, что и центр гравитации выказывает настойчивое присутствие, как только мы начинаем играть с ним. Однако никто никогда не видел и никогда не увидит центр гравитации. И как заметил Дэвид Юм, самости тоже никто никогда не видел.

«С моей стороны, когда я более всего приближаюсь к тому, что я называю моей самостью (myself), я всегда натываюсь на то или иное специфическое восприятие — жары или холода, света или тени, любви или ненависти, боли или удовольствия. Я никогда не могу схватить себя самого (myself) в какой-то момент без восприятия и никогда не могу наблюдать что-либо помимо восприятия... Если в результате серьезной и непредвзятой рефлексии кто-нибудь подумает, что у него име-

446

ется другое понятие его самого (himself), я с сожалением должен констатировать, что я больше не могу вести с ним разумный разговор. Все, что я могу позволить ему, — это признать, что он может быть таким же правым, как и я, но что мы существенно различны именно в этом. Он может, вероятно, мыслить что-нибудь простое и длящееся, что он называет себя самим (himself), хотя я уверен, что во мне нет такого принципа»<sup>1</sup>.

### 3. Условия личностного

Dennett D. Conditions of Personhood // D.C. Dennett. Brainstorms. Philosophical Essays on Mind and Psychology. Ch. 14. MIT Press, Cambridge Mass., L., 1986. P. 267—299. [Third printing, copyright in

1978. Материал 14 главы впервые опубликован в: The Identities of Persons. Ed. by A. Rorty. Un. of California Press, 1976].

Я — личность и вы тоже личность. Этот тезис не вызывает никаких сомнений. Я — человеческое бытие и, *вероятно*, вы также являетесь им. Если вы обиделись на слово «вероятно», вас можно обвинить в некоем виде расизма, ибо важным является не принадлежность к одному и тому же биологическому виду, а то, что мы — личности. Относительно же последнего я не выразил сомнения. Достоинство кого-либо определяется не его биологическим происхождением, не тем, что он рожден женщиной и родился вообще. Обычно мы игнорируем это обстоятельство, усматривая в человеческом определяющий признак личностного. Происходит это, несомненно, в силу коэкстенсивности или почти коэкстенсивности данных терминов в нашем контексте. В месте и времени, в котором

<sup>1</sup> Hume D. Treatise on Human Nature, 1. IV. Sec. 6.

447

мы живем, человеческие существа признаны единственными личностями, и почти всех их мы считаем личностями. Хотя нетрудно вообразить биологически весьма иных личностей, — населяющих, скажем, другие планеты. Мы также признаем наличие условий, освобождающих человеческие существа от личностного или, по меньшей мере, от некоторых очень важных элементов личностного. Например, новорожденным человеческим существам, существам с ментальными дефектами, существам, признаваемым профессиональными психиатрами нездоровыми, — им отказано в личностном или, во всяком случае, в существенных элементах личностного.

Казалось бы, что у столь важного понятия, как личность, которое мы с уверенностью прилагаем и отнимаем, должны быть четко сформулированные необходимые и достаточные условия правомерного его применения, однако, если они и есть, их еще нужно найти. Возможно, что в конце концов окажется, что искать нечего, а полученный вывод — понятие личности непоследовательно и устарело. Такого рода предположение высказал Б. Скиннер, однако оно не встретило понимания частично из-за трудности и даже невозможности представить состояние существа, если мы отбросим понятие личностного. Можно доказать (см.: «Brainstorms», ch. 12) самопротиворечивость идеи о возможности отказа видеть в других людях и *в нас самих* личностей (если только это не означает просто самоуничтожение и соответственно прекращение мысли о чем-то как о чем-то). Поэтому безотносительно к наличию правильного или ложного в основаниях, приведших Скиннера к выводу об устарелости понятия личности, трудно представить его победу в конкуренции с этим интуитивно неуязвимым представлением. Тем не менее, если даже понятие «личность» каким-то образом входит в незелиминируемую часть нашей концептуальной схемы, оно может пребывать в ней все еще в весьма путанном значении, нежели нам того хотелось. Может оказаться, скажем, что понятие личности — только свободно циркулирующее почетное звание, с удовольствием прилагаемое к себе самим и к другим людям по мере того, как нами

448

движет дух, подстегиваемый нашими эмоциями, эстетическими чувствами, соображениями политики и тому подобным, — точно так же, как понятие «модный», например, относится ко всем и только тем людям, которые позволяют рассматривать себя модными другим людям, тоже считающих себя модными. Действительно, быть личностью оказывается *чем-то* вроде этого, а если в бытии личности нет ничего большего, нам надлежит, если удастся это сделать, пересмотреть важную значимость, приписываемую этому понятию.

Предполагая в бытии личности наличие чего-то большего, ищущий необходимые и достаточные условия для обоснования этого большего все же может столкнуться с трудностями в случае наличия более чем одного понятия личности, а для данного подозрения есть основания. Короче, видимо, мы имеем дело с двумя переплетенными понятиями, одно из которых можно назвать моральным понятием, другое — метафизическим понятием личности. Локк говорит, что

«личность» — «это судебный термин, охватывающий и действия, и их оценку, поэтому он прилагается только к мыслящим агентам, способным подчиняться закону, быть счастливыми или несчастными. Личность сознательно простирает себя за настоящее в прошлое и таким образом становится заинтересованной и вменяемой» (Locke J. *Essays*, Book II, ch. XXVII). *Совпадает ли* метафизическое понятие — в простом смысле как интеллектуального, сознающего, чувствующего агента — с моральным понятием, в смысле ответственного, обладающего и правами и обязанностями агента? Или это означает только, что бытие личностью в метафизическом смысле — необходимое, но недостаточное условие для бытия личности в моральном смысле? Тожественно ли бытие сущности, которой приписываются состояния сознания и самосознания, с бытием как самоцелью, или же первое — только одно из предварительных условий? Следует ли считать деривацию из оригинальной позиции в теории справедливости Джона Роулса демонстрацией того, как метафизические личности *могут* становиться моральными личностями, или деривацию следует понимать в смыс-

449

ле демонстрации причин, почему метафизические личности *должны быть* моральными личностями<sup>1</sup>. Данное различие, выраженное на менее техническом языке, выглядит достаточно ясным: объявляя человека нездоровым, мы перестаем относиться к нему как к ответственному существу и не признаем за ним большинства прав, хотя наше общение с ним на практике не отличается от нормальных личностных взаимодействий. У понятия «личность» есть один смысл, позволяющий прилагать его к нездоровому человеку и видеть в нем «личность». В самом начале я заявил, что и вы и я являемся личностями, и это бесспорное положение. У меня не было оснований предполагать, а тем более приводить доказательства, что все читатели этого очерка юридически здоровы и морально ответственны. То, о чем можно было бы говорить, — вне всякого сомнения, если вообще можно говорить, что это бытие, к которому обращаются с личностными местоимениями «вы» и «я», было бы личностью в метафизическом смысле. Однако если это единственное не вызывающее сомнения по-

<sup>1</sup> В работе «Justice as Reciprocity» — переработанной версии статьи «Justice as Fairness» — Д. Роулз предлагает включить в понятие личности в оригинальной позиции «нации, провинции, деловые фирмы, церкви, команды и т.п. Принцип справедливости применим к конфликтующим претензиям, проявляемым личностями во всех этих разных сферах. Возможно, стоит говорить о некотором логическом приоритете, когда речь идет о человеческих индивидах» (Utilitarianism. ed. Gorovitz S. Indianapolis, 1971. P. 245). В книге «Теория справедливости» (Rawls J. *Theory of Justice*. Cambridge Mass., 1971) Роулз допускает, что партии в оригинальной позиции могут включать в себя ассоциации и другие группы, не являющиеся человеческими существами (p. 146), а очевидная взаимозаменяемость выражений «партии в оригинальной позиции» и «личности в оригинальной позиции» предполагает его утверждение, что для определенного морального понятия личности последняя *составлена* из метафизических личностей, которые сами могут быть, а могут и не быть моральными личностями.

450

ложение, тогда метафизическое и моральное понятия личности следует разделять. И все же предполагая различие этих понятий, представляется, что имеются все основания полагать метафизически личностное необходимым условием морально личностного<sup>1</sup>.

Сейчас я предлагаю рассмотреть шесть знакомых тем, каждая из которых, я полагаю, представляет собой заявку на определение необходимых условий личностного или некоторую интерпретацию. В качестве предмета обсуждения выбраны три вопроса: первый — каким образом (с точки зрения моей интерпретации) эти темы взаимозависимы; второй — почему они являются необходимыми условиями морально личностного; третий — почему так трудно сказать, составляют ли они все вместе достаточные условия морально личностного. *Первая* и наиболее знакомая тема — личности являются *рациональными существами*. Ее можно найти и в этических теориях Канта и Роулза, и в «метафизических» теориях Аристотеля и Я. Хинтикки<sup>2</sup>. *Вторая* тема — личности являются существами, атрибутивно наделенными состояниями сознания или такими, кому приписываются психологические, или ментальные, или *интенциональные предикаты*. Согласно Питеру Стросону,



например, личность есть «понятие некоторого типа сущности, к которой применимы *оба* предиката — и предикат, приписывающий состояние сознания, и предикат, приписывающий телесные характеристики»<sup>3</sup>. *Третья* тема — зависит ли

<sup>1</sup> В этой работе мы не касаемся толкования Роулсом сложных личностей. Более подробно об этом говорит Эмили Рорти (Rorty A. *Persons Policies and Bodies* // *International Philosophical Quarterly*. XII. 1. March 1973).

<sup>2</sup> Hintikka J. *Knowledge and Belief*. Ithaca, 1962.

<sup>3</sup> Strawson P.F. *Individuals*. L., 1959. P. 101—102. Определение Стросона часто критикуют за чрезмерную широту: оно охватывает все чувствующие активные существа. См., например: Frankfurt H. *Freedom of the Will and the Concept of a Person* // *The Journal of Philosophy*. LXVIII. Jan. 14, 1971. Можно показать (что я и собираюсь сделать),

451

каким-то образом то, что считается личностью, от *принятого к ней отношения*, от усвоенной особой установки (stance)<sup>1</sup>. Данная тема предполагает, что суть дела состоит не просто в том, что, однажды установив факт существования бытия чего-то как личности, мы относимся к этому чему-то соответствующим образом, а в том, что само наше обращение с ним (с ней, с этим) данным специфическим образом в какой-то форме и в какой-то мере конституитивно для бытия этого чего-то. Вариации на эту тему можно найти в работах Д. МакКея, П. Стросона, Р. Рорти, Х. Патнэма, У. Селларса, Э. Флю, Т. Нагеля, Д. Ван де Вейта, а также в моих работах<sup>2</sup>. *Четвертая* тема — объект, по отношению к которому принята такая личностная установка, каким-то образом способен отвечать взаимностью (reciprocating). Данную тему или ее наметки можно найти в работах Дж. Роулза, Д. МакКея, П. Стросона, Х. Грайса и других авторов. Взаимность (reciprocity) порою выражают

что состояния сознания представляют собой только подсистему психологически или интенционально характеризуемых состояний, а Стросон, как мне представляется, в данном случае явно нацелен достаточно широко раскинуть сети для охвата ими всех возможных психологических состояний.

<sup>1</sup> Английское слово «stance» (позиция, специфический подход, особое видение) в контексте данной работы Деннета мы переводим как «особая установка» (*прим. перев.*).

<sup>2</sup> MacKay D. *The Use of Behavioral Language to Refer to Mechanical Processes* // *British Journal of Philosophy of Science*, 1962. P. 89—103; Strawson P. *Freedom and Resentment* // *Studies in the Philosophy of Thought and Action*. Ed. by Strawson P., Oxford, 1968; Rorty A. *Slaves and Machines* // *Analysis*, 1962; Putnam H. *Robots: Machines or Artificially Created Life?* // *The Journal of Philosophy*, LXI, Nov., 1964; Sellars W. *Fatalism and Determinism* // *Freedom and Determinism*. Ed. by Lehrer K., N.Y., 1966; Flew A. *A Rational Animal* // *Brain and Mind*. Ed. by Smythies J., L., 1968; Nagel T. *War and Massacre* // *Philosophy and Public Affairs*, Winter 1972; Van de Vate D. *The Problem of Robot Consciousness* // *Philosophy and Phenomenological Research*, Dec. 1971.

452

посредством весьма неинформативного лозунга: быть личностью, значит, относиться к другим как к личностям. Из него делается вывод, возможно, следующий золотому правилу, что видеть в другом личность, значит, считать его моральным, но в таком случае смешиваются различные виды взаимности. Согласно Т. Нагелю, «крайне враждебное отношение к другому совместимо с личностным к нему отношением» (с. 134). Ван де Вейт обращает внимание на одно из различий, имеющих между некоторыми формами непредумышленного и умышленного убийства; в последнем случае убийца относится к своей жертве как к личности.

*Пятая* тема — личности должны обладать способностью к *вербальной коммуникации*. Это условие легко выводит животное не-человека из объема понятия личностного и, соответственно, освобождает его от моральной ответственности: данное условие, во всяком случае имплицитно, присутствует во всех социальных контрактных теориях этики. Тема вербальной коммуникации акцентировалась многими, работающими в философии сознания, включая меня самого, когда моральное измерение личности не выделялось в специальный предмет рассмотрения. *Шестая* — личности отличаются от других существ особым типом *осознания* (conscious): этот тип свойствен нам и не свойствен другим видам существ; порою его определяют как ту или иную форму *само-*

сознания. В работах трех современных философов — Э.Д. Энскомб («Интенции»), Ж.П. Сартра («Трансцендентальность Эго») и Х. Франкфурта (недавняя его статья «Свобода воли и понятие личности») — с разной аргументацией утверждается, что специфический тип сознательности (самосознание — *Н. Ю.*) является предпосылкой бытия морального агента<sup>1</sup>.

Я буду доказывать, что порядок моего представления перечисленных шести тем соответствует порядку их зависимости

<sup>1</sup> Frankfurt H. Freedom of the Will and the Concept of a Person // The Journal of Philosophy, LXVIII, Jan. 1971.

453

(dependence). С одной оговоркой. Она состоит в том, что первые три темы внутренне взаимозависимы; быть рациональным, значит, быть интенциональным и быть объектом определенной установки. Вместе взятые, они составляют необходимые, но недостаточные условия для показа формы взаимности. Последняя, в свою очередь, является необходимым, но недостаточным условием обладания способностью к вербальной коммуникации, представляющей собой необходимое условие для наличия особого типа сознания, являющегося, по утверждению Элизабет Энскомб и Хэрри Франкфурта, необходимым условием морально личностного.

Я уже касался тем рациональности, интенциональности и установки в других моих работах, посвященных определению не личности, а более широкого класса того, что я называю *интенциональными системами*. Поскольку мои рассуждения будут строиться на этом последнем понятии, имеет смысл дать ему краткую характеристику. Под интенциональной системой понимается система, поведение которой может быть (по меньшей мере иногда) объяснено и предсказано, основываясь на приписывании ей *верований и желаний* (и других свойств с интенциональными характеристиками, которые я буду именовать *интенциями*, включая в последние надежды, страхи, восприятия, ожидания и т.п.). *В каждом случае* возможны иные способы предсказания и объяснения поведения интенциональных систем, например, механические или физические. Однако интенциональная установка может оказаться удобной и самой эффективной, во всяком случае некоей успешной установкой, которую имеет смысл принять для представления объекта в виде интенциональной системы. При таком определении очевидно, что не все интенциональные системы являются личностями. Мы приписываем верования и желания собакам и рыбам и таким образом предсказываем их поведение. Можно даже прибегнуть к данной процедуре для предсказания поведения машин. Например, имея в качестве противника хорошо играющий в шахматы компьютер, приписывание ему интенций является хорошей, а на практике единственно хоро-

454

шей стратегией. *Допустив* наличие у компьютера определенных верований (или информации) и желаний (или функций предпочтения) успешно провести шахматную игру, я могу при благоприятных условиях вычислить наиболее вероятный следующий шаг компьютера в *случае допущения посылки о рациональном использовании компьютером верований и желаний*. В данном примере компьютер выступает в качестве интенциональной системы не потому, что обладает какими-то внутренними свойствами или что ему реально и истинно присущи верования и желания (какие бы они ни были). Эта роль обретается им в силу подчинения определенной *установке*, принятой по отношению к нему, а именно — интенциональной установке. Она реализуется путем приписывания интенциональных предикатов, при обычных ограничениях по отношению к компьютеру, и осуществляется при условии рассмотрения компьютера в качестве рациональной, на практике разумно рассуждающей сущности.

Важно иметь в виду ограниченность такого определения *интенциональной системы* и, соответственно, осознавать, что класс интенциональных систем может охватывать гораздо больше. Если я, положим, предскажу, что некоторое растение, например находящийся за углом дома плющ в горшке, будет тянуться вверх к свету в силу «стремления к свету» и «желания» выйти из-под закрывающей его тени, что оно «ожидает» или «надеется» найти за углом свет, я уже принял интенциональную установку по отношению к растению, и — о чудо! — в очень узких

пределах она работает. Коли она работает, некоторые растения представляют собой интенциональные системы, хотя и очень низкого уровня.

Практическая полезность принятия интенциональной установки по отношению к растениям стала близкой и понятной мне в беседе с лесорубами в лесах Мейна. Они без колебаний называли дерево не «это», но «он» и «она», о молодой ели говорили: «Она хочет распространить свои ветви, но не следует позволять ей это делать; ей захочется еще больше расширяться, чтобы получить свой свет». У них в ходу такие

455

фразы: «В отличие от кедров сосны не любят держать свои ноги в сырости». Вы можете «обмануть» яблоню, заставив ее «думать, что пришла весна»; для этого поздней осенью разведите костер под ее сучьями и она зацветет. Такой способ речи не есть просто образность и совсем не дань предрассудкам. Это эффективный способ осмысления, контроля, предсказания и объяснения поведения растений, дающий возможность легко обойти незнание контролирующих механизмов. Профессиональные биологи, возможно, предпочтут говорить об информационной трансмиссии от периферии дерева к другим его частям. Их язык будет менее красочным, тем не менее интенциональным. Полное воздержание от интенционального языка столь же героическое, громоздкое и бесцельное дело, как и накладывание бихевиористами жесткого табу на использование интенциональных понятий по отношению к крысам и голубям. И даже если интенциональное толкование той же активности дерева обладает малой эвристической ценностью, более мудро принять дерево за очень неразвитую, малоинтересную, незначительную интенциональную систему, нежели пытаться обозначить уровень, только выше которого интенциональные интерпретации признаются «объективно истинными».

Очевидно, что свойство быть интенциональной системой недостаточное, но, конечно, необходимое условие для бытия личности. Объект, к которому нам не удалось применить интенциональную установку, не может считаться личностью. Учитывая это, имеем ли мы право относить личности к подклассу интенциональных систем? На первый взгляд, предположение о личностях как подклассе интенциональных систем, *реально* имеющих верования и желания, а не просто имеющих их *предположительно* для краткосрочного их предсказания, может показаться плодотворным. Однако попытки сказать, что же следует считать реально имеющим верование (так, чтобы в определение не вошли ни собака, ни дерево, ни компьютер), по-видимому, упираются в выдвигание условий к истинному верованию, которые 1) являются слишком

456

для наших интуиций и 2) апеллируют к определенным условиям личностного, выходящим за рамки приведенного мною перечня. Например, кто-то может утверждать, что истинные верования с необходимостью *вербально выразимы* теми, кто разделяет их<sup>1</sup>, или что верящие в них должны *осознавать* их наличие. Однако люди, по-видимому, имеют множество верований, не поддающихся словесному выражению или не осознаваемых ими; во всяком случае, я надеюсь показать, что и способность к вербальному выражению, и способность к осознанию имеют различные loci в системе необходимых условий личностного.

Большой прогресс достигается, как мне представляется, при обращении к четвертой теме — *взаимности* — и рассмотрении возможных ее определений в терминах интенциональных систем. Данная тема предполагает способность личности отвечать взаимностью на установку, а именно такую, когда интенциональная система, принимающая интенциональную установку относительно других объектов, сама отвечает этим требованиям. Давайте определим *интенциональную систему второго порядка* как такую систему, которой мы приписываем не только простые верования, желания и другие интенции, но и верования и другие интенции *относительно* верований, желаний и других интенций. Интенциональная система «S» будет интенциональной системой второго порядка, если среди приписываемых ей свойств есть такие, как: *S верит, что T желает, что p, S надеется, что T боится, что g, и есть рефлексивные случаи вроде: S верит, что S желает, что p.* (Важность рефлексивных случаев будет особенно подчеркиваться нашим дальнейшим

обращением к авторам, интерпретирующим наше шестое условие как само-сознание. Может показаться, что рефлексивные случаи автоматически превращают все интенциональные системы в системы второго порядка и даже в системы n-порядка на том основании, что ве-

<sup>1</sup> Williams B. *Deciding to Believe // Language, Belief and Metaphysics*. Ed. by Kiefer H., N.Y., 1970.

457

рование, что *p* включает в себя верование, что вы верите, что *p* и т.д., однако в этом случае допускается существенная ошибка; повторение верований и других интенций никогда не является излишним, и отсюда, в то время как некоторые повторения нормальны и ожидаемы, они никогда не тривиальны или автоматичны.)

Следующий вопрос эмпирический: исходя из уже известного нам, можно ли сказать, что человеческие существа являются единственными интенциональными системами второго порядка? Мы, например, приписываем верования и желания собакам, кошкам, львам, птицам, дельфинам и в нормальной ситуации на практике часто предсказываем их поведение, но в случае очень софистичного поведения животного возникают трудности. Для объяснения или предсказания его поведения появляется желание приписывать ему интенции второго порядка. Конечно, если некоторая версия физикализма истинна (а я верю в это), у нас нет никакой нужды приписывать кому-либо какие-либо интенции. Но предположив, что для эвристических и прагматических целей нам все же следовало бы приписывать животным интенции, будем ли мы чувствовать прагматическое сопротивление, приписывая им интенции второго порядка? Психологи часто прибегают к такому принципу — своего рода бритве Оккама: для объяснения поведения организма ему следует по возможности меньше приписывать интеллекта, или сознания, или рациональности. Этот принцип может и фактически интерпретируется в пользу прямолинейного радикального бихевиоризма<sup>1</sup>, но я считаю это ошибкой; смысл его может быть интерпретирован как требование при принятии интенциональной установки по отношению к объекту приписывать ему простейшие наименее софистичные верования, желания и т.д. самого низшего уровня, объясняющие его поведение. В этом случае мы будем допускать, например, что пес Фидо *хочет* получить свой ужин и *верит*,

<sup>1</sup> Skinner B. *Behaviorism at Fifty // Behaviorism and Phenomenology*. Ed. by Wann T. Chicago, 1964.

458

что хозяин даст его, если он встанет перед хозяином и изобразит желание. И все же нам не следует приписывать Фидо тип *верования*, будто его просьба вызывает у хозяина *веру*, что он, Фидо, хочет получить ужин. Точно так же как из моего *ожидания*, после того как я положил монету в прорезь продающей сладости машины, не следует верование, что вложенная монета побуждает машину *верить* в мое желание получить конфеты. То есть хотя просьбы Фидо выглядят весьма похожими на подлинное взаимодействие второго порядка (Фидо относится к хозяину как к интенциональной системе), все же, предположив, что для Фидо хозяин не что иное, как машина, выдающая ужин, активизируемая с помощью просьб, мы будем иметь лучшее предсказательное приписывание свойств; более скромное, но все же, конечно, интенциональное.

Способны ли собаки, шимпанзе или другие «высшие» животные подняться до уровня интенциональных систем второго порядка, а если нет, то почему? Я привык думать, что не способны, исходя из посылки об отсутствии языка у животных, без чего невозможна репрезентация интенций второго порядка. Иными словами, я полагал, что условие четвертое основывается на условии пятом. Моя гипотеза состояла в том, что животные не могут иметь верования второго порядка — верования о верованиях — по той же самой причине, по которой они не могут иметь верования о пятнице или поэзии. Некоторые верования могут приобретаться опосредованно и, следовательно, репрезентироваться только языком<sup>1</sup>. Однако из признания истинным положение о невозможности приобретения некоторых верований без языка не следует истинность утверждения, что оно охватывает все верования второго порядка. Тезис, отрицающий возможность не-человеков быть интенциональными системами второго порядка, является лож-

<sup>1</sup> Для прояснения предположения об отношении языка к верованию и рациональности см.: Sousa R. *How to Give a Piece of Your Mind // Review of Metaphysics*, Sept 1971.



459

ным. Как-то я обратился с просьбой к моим знакомым привести примеры не-человеческих интенциональных систем второго порядка, и мне рассказали о нескольких весьма убедительных случаях. Рассмотрим один из них, приведенный в письме Питера Эшли.

«Однажды вечером я сидел у себя дома в кресле, единственном из всех, в котором моей собаке разрешалось спать. Собака лежала передо мной и скулила. Прибегнув к разным и безуспешным попыткам «убедить» меня освободить кресло, она предприняла следующее действие — и оно было самым интересным или даже, я бы сказал, единственно интересным моментом всей этой истории. Она встала и пошла к выходу, не выходя из моего поля зрения. Поцарапала дверь, и у меня создалось впечатление, что она отказалась от попыток захватить кресло и решила прогуляться. Но стоило мне подойти к двери, как она ринулась обратно в комнату и взобралась на любимое кресло, которое она "вынудила" меня покинуть».

Казалось бы, здесь мы должны приписать собаке интенцию (что хозяин поверит в ее желание выйти на улицу) не только второго, но и третьего порядка. Ключевым в этом примере, делающем его случаем интенциональной системы высшего порядка в действии, является факт, что верование, которое собака намеревалась внушить хозяину, ложное. Если мы пожелаем найти другие примеры поведения животных — интенциональных систем второго порядка, полезно рассмотреть случаи обмана, когда животное, веря в *P*, пытается заставить другие интенциональные системы верить в не-*P*. Там, где животное стремится вызвать у другого поведение, к которому его не принуждает его истинное верование относительно положения дел, мы не можем «полностью отгораживаться» и прибегать к объяснениям, апеллирующим только к интенциям первого уровня. Перед тем как перейти к объяснению, почему это так, данный тезис можно сделать более общим: если *X* пытается внушить *Y* поведение, неадекватное его *истинным* окружающим условиям и потребностям, но соответствующее *воображаемым* положению дел, *верованиям* и по-

460

требностям, мы вынуждены приписывать *X* интенции второго порядка. В этой форме тезис выглядит знакомым — к нему часто прибегают критики бихевиоризма: оставаться бихевиористом, говорят они, можно объясняя и контролируя поведение лабораторных животных, если руководствоваться положением об отсутствии серьезного расхождения между действительными условиями эксперимента и условиями, воспринимаемыми самими животными. Лабораторная тактика испытывающих затруднения бихевиористов состоит в проведении экспериментов, обманывающих субъектов опыта: если обман достигает цели, поведение субъектов предсказуемо с точки зрения их ложных *верований* относительно положения дел, а не с точки зрения их действительного состояния. В этом случае интенциональная система первого уровня является неким бихевиористом; она ничему не приписывает интенции. Поэтому, если нам нужно иметь веское опытное свидетельство того, что интенциональная система *S* не является бихевиористской, то есть относится к интенциональной системе второго порядка, мы его получаем только в тех случаях, когда бихевиористские теории неадекватны данным опытам и бихевиоризм не в состоянии объяснить успех системы *S* в манипулировании поведением другой системы.

Все это наводит на мысль, что пример Эшли в конечном итоге не столь убедителен и может быть опровергнут предположением, что его собака является бихевиористом того или иного толка. Ей не нужно было верить, скребясь в дверь, что она внушит Эшли мысль о своем желании выйти из дома; как хороший бихевиорист она просто может верить, что, скребясь в дверь, она обусловит Эшли пойти к двери. Иначе говоря, она применяет обычный стимул, получает обычную реакцию и кроме этого здесь нет ничего другого. Случай с Эшли поддается объяснению, если это *обычный* способ, каким его собака добивается открывания двери. Вероятно, это так и есть, тогда более скромная гипотеза состоит в наличии у собаки верования, что ее хозяин, видя как она скребется, обусловлен пойти к двери. Если бы собака предприняла что-

461

либо *новое* для обмана своего хозяина (как-то: подбежала к окну и, рыча подозрительно, выглянула из него), тогда мы должны были бы допустить, что вставание с кресла не было просто



условной реакцией Эшли и как таковое не было «видно» собаке, однако допущение такой виртуозности у собаки в высшей степени неправдоподобно.

И все же в чем разница между неправдоподобным случаем обмана и хорошо подтвержденными случаями, когда бегущая по земле птица, стремясь увести преследователя от гнезда, делает вид, будто у нее сломано крыло. Достигнутый здесь эффект является новым в том смысле, что птица многими повторениями не вырабатывала у местных преследователей такой стимул, и вроде бы мы вынуждены объяснять эту уловку в чем-то как подлинный обман с *интенцией* побудить у преследователя ложное *верование*. Подталкиваемые к такой интерпретации, мы крайне поразились бы изобретательности птицы, не будь нам известно о «просто инстинктивном» поведении такого рода. Но почему нужно считать, что обозначение такого трюка «просто инстинктивным» есть его умаление? Утверждать его инстинктивность — значит утверждать, что все птицы данного вида проделывают этот трюк: они проделывают его даже в неподходящих условиях, прибегая к нему, когда разумнее оставаться в гнезде. Их модель поведения жесткая — своего рода тропизм, — контроль, вероятно, генетически встроен, а не усвоен или изобретен.

Говоря об инстинктивном, мы должны быть осторожными и не заходить в его умалении слишком далеко; дело не в «бездумности» проделываемого птицей трюка; хотя не вызывает сомнения, что она проделывает его, не прибегая к аргументу или плану в ее голове (типа «Попробую сделать так: если я свешу крыло, как будто оно сломано, лиса может подумать, что...»). Человек тоже может делать что-нибудь схожее по тонкости, подлинной интеллектуальности, новизне и соответствии обстоятельствам, вовсе не проводя умственные шаги через «осознанные мысли». *Мышление о мыслях*, как бы его ни определяли, не делает подлинно интеллектуальное поведение ин-

462

теллектуальным. Энскомб говорит по этому поводу: «Предположение о том, что [такое выражение разумного рассуждения] служит для описания ментальных процессов, было бы в целом совершенно абсурдным. Объяснение нацелено на описание порядка, который наличествует всякий раз, когда действия совершаются с интенциями»<sup>1</sup>. Однако «порядок наличествует» и в случае с птицей, и в случае с человеком. То есть, задаваясь вопросом о причине наличия в эволюции птицы такого тропизма, мы усматриваем эту причину в полезности обладания средствами *обмана* преследователей или внушения у них ложных верований: на самом деле следует объяснять происхождение у птиц интенций второго порядка. С моей стороны было бы большим радикализмом отрицать или умалять существование огромных различий между инстинктивным или тропическим поведением птицы и более гибким и интеллектуальным поведением человеческих и иных существ, однако в данном случае я отстаиваю следующее положение: если принять без оговорок интенциональную установку за инструмент предсказания и объяснения, тогда птицу, как и любого человека, в той же мере следует считать интенциональной системой второго порядка. А коли это так, нам нужно весьма осторожно относиться к приводимому мною ранее аргументу, а именно, что репрезентации интенций второго порядка каким-то образом зависят от языка<sup>2</sup>. Вовсе не очевидно, нужно ли, чтобы все или даже некоторые из верований и интенций интенциональной системы каким-то образом были *репрезентированы* «внутри» этой системы, чтобы мы обрели преимущества в предсказании ее поведения, приписывая системе такие интенции (более детальное обоснование этого положения дано в «Brainstorms», ch. 2, 3). Ситуация, которую мы проясняли на примере стремления птицы внушить преследователю ложное верование, по-видимому, не имеет пространства или необходимости для репрезентации данной со-

<sup>1</sup> Anscombe G.E. *Intention*. Oxford, 1957. P. 80.

<sup>2</sup> Sousa R. *Self-Deception*. // *Inquiry*. XIII. 1970. P. 317.

463

фистичной интенции в «мыслях» или «сознании» любого существа, ибо ни птица, ни эволюционная история, ни Мать-Природа не имеют нужды мыслить эти мысли для подкрепления нашего объяснения.

В этом случае следует согласиться, что взаимность (reciprocity) — в смысле простой способности интенциональных систем демонстрировать интенции высшего порядка — зависит от первых трех

условий и не зависит от пятого или шестого. Насколько такое понимание справедливо по отношению к идее взаимности, обсуждаемой другими авторами, станет видно только при условии его стыковки с двумя последними условиями. Для рассмотрения пятого условия — способности к вербальной коммуникации — мы обратимся к теории значения Х. Грайса. Грайс пытается определить то, что он называет «ненатуральным» значением — придание произносящим значения чему-то путем высказывания об этом чем-то в терминах интенций произносящего. Его первоначальная дефиниция выглядит следующим образом<sup>1</sup>:

«*U* имеет в виду что-то, произнося *X*», и это является истинным, если для *A* (аудитории слушателей) *U* произнес *X* с намерением, чтобы:

1. *A* отреагировала на произносимое определенным откликом *R*.
2. *A* думала (признала), что *U* имеет намерение (1).
3. *A* выполнила (1) на основе выполнения (2). Заметьте, что интенция (2) приписывает *U* интенцию не

только второго, но и третьего порядка. *U* должен иметь намерение, что *A* распознает намерение *U*, чтобы у *A* была реакция *R*. То обстоятельство, что серия контрпримеров вынужда-

<sup>1</sup> Ключевыми для позиции Х. Грайса являются две его статьи: Grice H. Meaning // Philosophical Review, July, 1957; Grice H. Utterer's Meaning and Intentions // Philosophical Review, April, 1969. В ряде пунктов исходная формулировка, приведенная в первой статье, была подвергнута ревизии во второй. Данная формулировка взята из второй статьи (р. 151).

464

ла Грайса двигаться от этой первоначальной дефиниции все к более и более сложным версиям, не имеет существенного значения, ибо все они воспроизводят интенцию третьего порядка (2). В грайсовском анализе ненатурального значения для нас особо важны два момента: первый — поскольку ненатуральное значение или значение чего-то через высказывание о чем-то должно быть характеристикой любой подлинно вербальной коммуникации, а последнее зависит от интенций третьего порядка высказывающего, мы имеем наш случай, когда условие пятое основывается на условии четвертом, а не наоборот; второй — Грайс показывает, что для достижения подлинной взаимности недостаточно иметь просто интенции *второго* порядка, требуются еще интенции *третьего* порядка. Грайс вводит условие (2) для исключения случаев такого, например, вида: я оставил разбросанной фарфоровую посуду, разбитую моей дочерью, с тем чтобы это увидела моя жена. Это не случай придания значения чему-то путем делания того, что я делаю, или намерения предпринять, ибо хотя я пытаюсь таким образом заставить мою жену поверить во что-то относительно акта дочери (прибегая к интенции второго порядка), успех зависит не от узнавания ею моей интенции, или признания моего вмешательства, или вообще от моего существования. Между нами здесь не было реального столкновения (encounter) (удачный термин, используемый Эрвином Гофманом), не было взаимного признания (recognition). Для того чтобы произносящий сообщал какое-либо значение, он должен столкнуться, встретиться со слушателями, но это может происходить и в отсутствии не-натурального значения. (Наблюдение за собакой Эшли.) И уловкам, зависящим от интенций третьего порядка, нет необходимости иметь встречу-столкновение (например, *A* может иметь интенцию, что *B* верит, что *C* желает, что *p*). Поэтому интенции третьего порядка являются необходимыми, но недостаточными условиями для встреч-столкновений, а те, в свою очередь, — необходимыми, но недостаточными условиями для примеров ненатурального значения, т.е. примеров вербальной коммуникации.

465

Не случайно, что примеры ненатурального значения, приводимые Грайсом, попали в класс, другие члены которого — случаи обмана или манипуляции. Рассмотрим, например, удачный контрпример Джона Сёрля, направленный против одной из формулировок Грайса. В Италии во время Второй мировой войны один американец был захвачен итальянцами за линией фронта. Он пытался их обмануть, внушив мысль, что он является германским офицером, произнося единственно известную ему немецкую фразу: «Kennst du das Land, wo die Zitronen blühen?»<sup>1</sup>. Согласно Грайсу, и этот случай, и случаи ненатурального значения роднит апелляция или

эксплуатация рациональности жертвы. Успех зависит от возможности внушения жертве мысли о необходимости строить цепь рассуждений, в которую внушающий прямо или косвенно вводит посылки. При обмане сам обманщик не верит в посылки; при нормальной коммуникации в них обычно верят. Вид коммуникации, представленный Грайсом, выглядит разновидностью совместной манипуляции говорящего со слушателями; она зависит не только от рациональности аудитории, которой надлежит разбираться в интенциях произносящего, но и от доверия слушателей к говорящему. Даже при условии рациональности аудитории, коммуникация, как разновидность манипуляции, не заработает, пока доверие аудитории к говорящему не будет обоснованным или разумным. Таким образом, норма для высказывания — искренность, высказывания же, не имеющие обычной ценности доверия, не достигнут своей цели (см.: «Brainstorms», ch. 1).

Ложь как форма обмана может работать только при наличии правдивых высказываний, другие же формы обмана не зависят от доверия жертвы. В этих случаях успех определяется недостаточно высоким уровнем изощренности жертвы. Глупые игроки в покер, неспособные видеть предлагаемые им

<sup>1</sup> Searle J. What is a Speech Act? *Philosophy in America*. Ed by Black M., L., 1965. Его позиция обсуждается Грайсом в «Utterer's Meaning and Intention». P. 160.

466

блеф и уловки, — проклятие для умных игроков. Софистичные обманы не обязательно предполагают непосредственное столкновение сторон. Есть книга, описывающая способы выявления подделок антиквариата (понятно, что одновременно она может служить и пособием по созданию фальшивого антиквариата), в которой желающим одурачить покупателей, мнящих себя «экспертами», предлагается такой хитрый совет: сделав стол или какую-нибудь другую вещь (с использованием всех средств симуляции века и изношенности), возьмите современную электродрель и просверлите дырку в каком-нибудь скрытом и вызывающем недоумение месте. Предполагаемый покупатель будет рассуждать примерно так: никто не будет сверлить столь бесформенную дыру, не имея на то оснований (в любом случае нельзя предположить, что она выглядит «аутентичной»), поэтому у нее должна быть какая-то цель. Последнее наводит на мысль, что стол использовался в чьем-либо доме; а поскольку он был в употреблении, вряд ли его сделали специально для продажи в этой лавке антиквариата, ...поэтому стол — подлинный антиквариат. Даже если в цепи рассуждений остается место для смутных сомнений, покупатель столь занят мыслью о цели использования дыры, что потребуются месяцы, прежде чем эти глубинные сомнения всплывут на поверхность.

В приведенных примерах обмана, как и в случае птицы со свесившимся крылом, важен факт, что успех здесь не зависит от *сознательного построения* жертвой обмана цепи разумных рассуждений. Не имеет значения, если покупатель только заметил дыру и «интуитивно почувствовал» — вещь подлинная. Позднее он, *может быть*, примет основание, послужившее «веским доводом» в пользу «подлинности» вещи, но он может отрицать его, а отрицая, может обманывать себя, даже если подобные *мысли* и не мелькали в его голове. Цепь рассуждения объясняет, почему данная дырка сработала именно так (если она сработала), однако, как говорит Энскомб, нет необходимости «описывать действительные ментальные процессы», если мы считаем эти процессы сознательными процессами или

467

событиями. Сказанное безусловно верно и в отношении коммуникаций Грайса: ни говорящему, ни слушателям нет нужды сознательно проводить через себя описываемые Грайсом сложные интенции. (В связи с этим удивительно, что этот факт никем не был использован для возражения ему.) Критики часто упрекают Грайса за ограниченность предложенных им условий для значения, чтобы посчитать их за достаточные, но мне представляется, что есть еще один, до этого не замеченный аргумент, показывающий, что эти условия не являются даже необходимыми. Действительно, мало кто осознавал такие изобретательные интенции, пока Грайс не указал на них, тем не менее люди всегда вступали в коммуникацию. До Грайса, если бы кого-нибудь спросили:

«Хочешь ли ты, чтобы слушающий распознал твою интенцию спровоцировать некую реакцию у него?», ответ скорее всего был бы таким: «У меня не было подобных хитроумных намерений. Я просто хотел сообщить, что скорее всего не приду домой к ужину (или что-нибудь вроде этого)». Поэтому, если принять усложненные интенции за скрытую основу наших коммуникаций, они должны быть подсознательными интенциями. Вполне естественной реакцией на статьи Грайса было бы заметить, что человек, находясь в коммуникации, *не осознает*, какого рода ходы он проделывает. Сейчас Энскомб настаивает — и весьма решительно, — что такого рода реакция устанавливает: описываемое таким образом действие не было *интенциональным*<sup>1</sup>. Если в процессе разговора человек не *осознает* эти *интенции*, он не может говорить с такими *интенциями*.

Почему никто не использовал этот аргумент против теории Грайса? Полагаю, из-за того, что Грайс что-то нашупал, сформулировал необходимые условия для ненатуральных значений. Прделанный им анализ проясняет многое. Вступаем

<sup>1</sup> См.: Anscombe G.E. Intention. Oxford, Blackwells, 1957. P. 11. Здесь и в следующих нескольких параграфах я использую понятия «интенции» и «интенциональный» в их обычном смысле и в написании выделяю их курсивом.

468

ли мы в коммуникацию с компьютерами на языке Фортран? Фортран принят за язык, имеющий словарь, грамматику, семантику. Взаимодействие между человеком и компьютером на этом языке часто приводят в качестве примера *коммуникации человека с машиной*, однако такое взаимодействие — бледная копия человеческой вербальной коммуникации именно потому, что в ней упущены грайсовы условия для ненатуральных значений. Здесь нет пространства для их применения. Достижение определенных целей в передаче машине какой-то части языка Фортран не зависит от придания машине способности распознавать чьи-либо *интенции*. Это вовсе не означает, что в будущем все коммуникации с компьютерами сохранят эту ограниченность (или силу — все зависит от ваших целей), это значит только то, что сегодня мы не вступаем в коммуникацию с компьютером в строгом (грайсовом) смысле слова<sup>1</sup>.

Если мы еще не готовы отбросить модель Грайса, хотя и фиксируем отсутствие такого рода *интенций* в нашем нормальном разговоре, нам следует загнать их в подполье и называть бессознательными или предсознательными *интенциями*. Они представляют собой *интенции*, присутствующие в «порядке, который наличествует» в коммуникативных отно-

<sup>1</sup> Ховард Фридман подсказал мне возможность использования такого аргумента: многие современные устройства Фортран «корректируют» вводимые оператором данные, расставляют знаки «плюс», ставят скобки и т.д. и выдают хорошо оформленные высказывания, т.е. отвечают критериям Грайса, поскольку в рамках очень ограниченной области они диагностируют *интенции* «говорящего» и действуют на основе этого диагноза. В связи с этим я хотел бы обратить внимание на следующее: во-первых, машины на сегодня могут диагностировать только то, что может быть названо синтаксическими интенциями оператора и, во-вторых, эти машины, по-видимому, не отвечают последующим, более разработанным определениям Грайса. Но я вовсе не хочу утверждать, что этого не сможет сделать никакой компьютер.

469

шениях людей, *интенции*, которые мы обычно не осознаем *интенциями*, являющимися предусловием вербальной коммуникации.

Мы зашли слишком далеко, не обращаясь к какому-либо виду сознания, ибо если и существует зависимость между сознанием или самосознанием и другими нашими условиями, это должно быть сознание, зависящее от других условий. Однако для показа последнего я должен сначала продемонстрировать определенную роль, которую первые пять условий сами по себе могут играть в этике. Такое предположение сделал, например, Джон Роулз в теории справедливости. Центральными в его теории являются два момента: конструирование идеализированной ситуации — «оригинальной позиции», в которую ставятся идеализированные личности, и выведение из этой идеализации первых принципов справедливости; из последних выводится и проясняется все остальное. Сейчас меня интересует не содержание этих принципов или достоверность их выведения, а суть применяемой им тактики. Роулз исходит из допущения, что некая совокупность



идеализированных личностей — рациональных и движимых само-интересом субъектов — просчитывает, принимая определенные ограничения, вероятные и возможные взаимодействующие последствия индивидуальных и антагонистических интересов. (Это требует от них конструирования интенций высшего порядка, например, верования в желания других людей, верования о верованиях других людей относительно их собственных желаний и т.п.) Роулз утверждает, что результатом этих просчитываний является оптимальное «решение», которое было бы разумно принять каждой движимой само-интересом личности как альтернативу гоббсовскому естественному состоянию. Решение состоит в достижении договоренности с другими личностями твердо следовать принципам справедливости, в общих чертах обрисованным Роулзом. Какой тип доказательства используется им в данном случае? По Роулзу, принятие данных принципов справедливости может рассматриваться как решение «игры высшего уровня» или как

470

«вопрос сделки». Оно аналогично и деривациям теории игр, и доказательствам эпистемической логики Я. Хинтикки<sup>1</sup>, и «демонстрации» того, что играющий в шахматы компьютер будет делать определенный ход, поскольку он наиболее рационален согласно заданной ему информации. Все зависит от принятия посылки об идеальных рациональных просчитывающих агентах, следовательно, их результаты внутренне нормативны. В деривациях из роулзовской оригинальной позиции я вижу продолжение и экстраполяции, присутствующие в более простых применениях интенциональной установки для понимания и контроля за поведением более простых сущностей (см.: «Brainstorms», ch. 1). Как истина и последовательность являются нормами верований, а искренность — нормой разговора, точно так же, если Роулз прав, справедливость является определяющей нормой межличностного общения. Но тогда следует принять и другую аналогию: подобно тому как частичным основанием для признания наличия у какого-либо существа верований или других интенций является наша способность мыслить его *рациональным* существом, основание считать существо личностью предполагает нашу способность полагать его подчиняющимся принципам справедливости. Используемый им здесь способ представления особого статуса понятия личности таков: не имея намерения защищать тезис о справедливости как о необходимом результате *человеческого* общения, Роулз, в конечном итоге, защищает тезис, что она — необходимый результат *личностного* взаимодействия. То есть понятие личности по своему внутреннему содержанию по необходимости нормативно и идеализировано; люди только в той мере личности, в какой в их деяниях и взаимодействиях проявляется справедливость. И снова можно видеть «порядок, который наличествует» в справедливом обществе, независимый от любых действительных фактов сознательного мышления. Существование справедливых практик и имплицитно присущее им «осознание» не зависят от того, пред-

<sup>1</sup> См.: Hintikka J. Knowledge and Belief. P. 38.

471

принимает ли вообще кто-либо сознательные и продуманные расчеты идеализированной оптимальной позиции; не зависит оно и от сознательной выработки обоюдного соглашения, и от сознательного принятия установки в отношении другого. Роулз пишет:

«Признание кого-то личностью требует определенного способа отношения к нему и действия; этот способ подразумевает различные *prima facie* обязанности. Признание в какой-то мере этих обязанностей и тем самым обладание элементами морали не связано ни с выбором, ни с интуитивными моральными качествами, ни с выражением чувств или отношений... это просто следование одной из форм поведения, демонстрирующей, что другие тоже признаются личностями!».

Важность предпринятой Роулзом попытки вывести справедливость из «оригинальной позиции» состоит, конечно, в том, что в то время как вывод признается *моральной* нормой, он не *выводится* как моральная норма. Мораль не предполагается в отношениях сторон, находящихся в оригинальной позиции. Однако это означает, что выведение нормы само по себе не дает нам какого-либо ответа на вопрос — когда и почему у нас есть право считать личностей *морально* ответственными за деривацию из этой нормы. Помочь нам в этом может Энскомб, вводя в то же



время наше шестое условие. *Если меня считают ответственным за некое действие* (какую-то часть моего поведения при определенном описании), у меня должно быть осознание этого действия при данном описании<sup>2</sup>. Почему? Потому что только при условии осознания этого дей-

<sup>1</sup> Rawls J. Justice as Reciprocity // Utilitarianism. Ed. by Gorovitz S. Indianapolis, 1971. P. 259.

<sup>2</sup> Меня могут считать ответственным за события или состояние дел, не проходившие через мое сознание, но некоторые из которых я был обязан сознавать; однако они относятся к *интенциональным* действиям. В этих условиях я ответствен за развитие событий в силу ответственности за предсказуемые последствия действий — включая акты оплошности, которые мне известны.

472

ствия я имел бы возможность высказаться и участвовать, находясь в привилегированной позиции, в игре вопросов-ответов, приводя основания для своих действий. (Если у меня нет привилегированной позиции отвечать на вопросы относительно моих действий, задавать их мне нет особого резона.) Для участия в этой игре важно следующее: только обладающие способностью участвовать в приведении разумных оснований могут быть склонены или, наоборот, разубеждены аргументами о характере тех или иных действий или отношений; если же кто-либо не в состоянии «прислушиваться к голосу разума» в каких-то вещах, его нельзя считать ответственным. Способность к вербальной коммуникации и к осознанию своих действий существенна для проявляющих готовность прислушиваться к аргументу или убеждению. Такое убеждение и взаимное приспособление интересов, достигаемое во взаимной эксплуатации рациональности, является характерной чертой оптимальной модели личностного взаимодействия.

Способность участвовать во взаимном убеждении выступает основанием еще одного условия личностного, на которое недавно обратил внимание Хэрри Франкфурт<sup>1</sup>. По Франкфурту, личности относятся к подклассу интенциональных систем, способных, по его выражению, к *«волюниям второго порядка»* (second-order volitions). На первый взгляд они выглядят схожими с классом интенций второго порядка, о которых я говорил, однако, как мы покажем далее, это не так.

«Помимо хотения, выбора и склонности делать то или другое люди могут также хотеть иметь (или не иметь) определенные желания и мотивы. Им присуща способность хотеть быть иными в своих предпочтениях, нежели они на самом деле есть... По-видимому, кроме человека ни одному животному не свойственна способность к рефлексивной самооцен-

<sup>1</sup> Frankfurt H. Freedom of the Will and the Concept of a Person // Journal of Philosophy. LXVIII. Jan. 14, 1971. Франкфурт не поясняет, является ли предложенное им условие морально личностного только необходимым или также и достаточным.

473

ке, проявляющейся в формировании желаний второго порядка» (р. 7).

Франкфурт приводит пример, когда личность может сказать о своем хотении иметь некое особое желание, даже не имея охоты сделать это желание действительным для себя, «своей волей». (Например, кто-то может возжелать героин с целью познать чувство желания героина, вовсе не стремясь сделать такое желание своим действительным желанием.) В более серьезных случаях человек желает иметь отсутствующее у него в настоящее время желание и желает сделать его своей волей. Эти случаи Франкфурт называет волениями второго порядка, и их наличие, считает он, «существенно для бытия личности» (р. 10). Я не буду здесь подробно рассматривать его доводы, скажу следующее: аргумент опирается на анализ различия между понятиями иметь свободу действия и иметь свободу воли. Согласно его анализу, человек имеет свободу воли только тогда, когда он может иметь ту волю, какую он хочет, и когда могут быть удовлетворены его воления второго порядка. Личности не всегда имеют свободу воли и при некоторых обстоятельствах могут быть ответственны за действия, осуществленные при отсутствии свободы воли, однако личность всегда должна быть «существом, для которого свобода его воли может стать проблемой» (р. 14), то есть существом, способным к формированию волений второго порядка, независимо от того, реализуются эти воления или нет. Франкфурт вводит замечательный термин «wanton» для обозначения тех, «кто имеет желания первого порядка... не имея волений второго порядка». (К волениям второго порядка Франкфурт, конечно, относит все *рефлексивные*

желания второго порядка.) Он утверждает, что наши интуиции подкрепляют мнение о том, что животные, не являющиеся человеком, а также маленькие дети и некоторые люди с ментальными дефектами, являются wantons, и мне трудно представить убедительные контрпримеры в отношении этих случаев. Действительно, можно посчитать за сильную сторону его теории утверждение, что человеческие существа —

474

единственные, кого мы признаем за личности, — отличаются в этом отношении от животных. Но что можно посчитать за специфическое в волениях второго порядка? Почему среди интенций высшего порядка они составляют особую провинцию личности? Я полагаю потому, что рефлексивная «само-оценка», о которой идет речь, по Франкфурту, является и должна быть подлинным само-сознанием. Оно достигается только в результате принятия в отношении *себя самого* установки, видя в себе не только коммуникатора, но и агента в смысле Энскомб — вопрошающего об основаниях и убеждающего с помощью доводов. Франкфурт замечает, что желания второго порядка — пустое понятие, пока невозможно действие с ними, а такого рода действие логически должно отличаться от действия, базирующегося на компоненте первого порядка. Действие на основе желания второго порядка или какое-то делание для обретения желания первого порядка представляет собой активность, направленную на самого себя как если кто-либо направлял ее в отношении другого; человек *обучает* себя, убеждает себя, предлагает аргументы, угрожает, подкупает себя в надежде принудить себя обрести желание первого порядка<sup>1</sup>. В этих случаях и установка по отношению к себе, и *доступ к себе* по существу те же самые, что и по отношению к другим. Человек должен *спрашивать себя*, каковы в действительности его собственные желания, мотивы, основания, и только если он в состоянии осознанно сказать о них и осознать свои желания, он

<sup>1</sup> Мое внимание обратили на факт, что собаки в конюшне очень часто занимаются мастурбацией, по-видимому, для *увеличения* их желания к копуляции. Однако такие случаи вряд ли стоит принимать в расчет; даже если предположить, что собака действует с желанием усилить желание, результат все же достигается неинтенциональным способом («чисто физиологическим»); собака не апеллирует к собственной рациональности и не использует ее для достижения своей цели. (Как если бы единственным способом действий личности на основе волений второго порядка было бы принятие таблетки, стояние на голове и т.п.)

475

обретает позицию, позволяющую заставить изменить себя<sup>1</sup>. Только в этом случае, как мне представляется, имеет место ситуация, когда «порядок, который наличествует» не может быть в наличии без его наличия в эпизодах сознательного мышления, в диалоге с самим собой<sup>2</sup>.

В заключение зададимся вопросом, почему мы не обрели позицию, позволяющую считать перечисленные выше необходимые условия морально личностного также и достаточными условиями? Просто потому, что понятие личности, что я и попытался показать, неизбежно нормативно. Человеческие или другие существа могут только стремиться приблизиться к этому идеалу, однако способа установления «ступеней приближения» к нему не существует, ибо он всегда будет произвольным. Если даже шесть условий (при точной интерпретации) посчитать за достаточные, они все равно не гарантируют, что какая-либо актуальная сущность

<sup>1</sup> Маргарет Джилберт исследует выводы из того факта, что «только когда кто-то верит, что ему присуща какая-то черта, он может принять решение освободиться от нее». Gilbert M. Vices and Self-Knowledge // Journal of Philosophy. LXVIII. Aug. 1971. P. 452.

<sup>2</sup> К. Маркс в «Немецкой идеологии» говорит: «Язык, как и сознание, возникает только из потребности, из необходимости, из взаимодействия с другими людьми... Язык столь же давен, как и сознание, язык — это практическое сознание». У Ф. Ницше в «Радостной мудрости» есть такие слова: «Поскольку является фактом наша способность думать, чувствовать, проявлять волю и обращаться к памяти, мы таким же образом способны «действовать» в полном смысле этого термина, и тем не менее ничто из всего этого не требует «прохождения через сознание», говоря метафорически»;.. Какова же тогда цель сознания вообще, когда оно в главном поверхностно? — Сейчас мне представляется, если вы выслушаете мой ответ и заключенное в нем, может быть, экстравагантное предположение, что тонкость и сила сознания всегда пропорциональны

способности к коммуникации в человеке (или животном), а способность к коммуникации, в свою очередь, пропорциональна необходимости в коммуникации. .. Короче, развитие речи и развитие сознания (не разума, но разума, становящегося самосознающим) идут рука об руку».

476

есть личность, ибо ничто никогда не отвечало бы этим условиям. Моральное понятие личности и метафизическое понятие личности не отдельные и различные понятия, а две различные и подвижные точки остановок в рамках одного и того же континуума. Эта релятивность сказывается на качестве удовлетворительности условий личностного на всех уровнях. Не существует объективно удовлетворительного, достаточного условия, позволяющего утверждать, что какое-то существо *в действительности* имеет верования. И по мере того как мы обнажаем очевидную иррациональность, лежащую за интенциональной интерпретацией существа, наши основания для приписывания вообще кому-либо верований испаряются. Особенно тогда, когда в наших руках находится (что, в принципе, мы всегда имеем) не-интенциональное, механистическое объяснение существа. Точно таким же образом допущение, что некое существо является личностью, оказывается шатким как раз в тех случаях, где оно особенно практически значимо — при совершении каких-то проступков и возникновении вопроса об ответственности. Ибо в этих случаях основания для признания личности виновной (свидетельства неправильных действий при осознании их неправильности и совершения их по своей собственной свободной воле) сами по себе выступают основаниями для сомнения: а является ли вообще тот, с кем мы имеем дело, личностью. И если бы меня спросили, что же может *разрешить* наши *сомнения*, ответ будет — ничего. При возникновении такого рода проблем мы не можем даже в отношении нас самих сказать, являемся ли мы личностями.

#### 4. Почему нам важно понимать это правильно. Постмодернизм и истина

Доклад на 19 Всемирном философском конгрессе. Бостон, 13 августа 1998 г. Daniel Dennett. Why Getting It Right Matters. Postmodernism and Truth // Free Inquiry. V. 20. № 1.

477

Winter 1999/2000. P. 40—43. Текст доклада представляет собой часть лекции, прочитанной в феврале 1997 года в Оксфорде. (Полный текст опубликован в «Values of Science». Ed. by W. Williams, Oxford, 1999. P. 95—109.)

Я расскажу историю, которую вы, вероятно, не слышали, о том, как группа американских исследователей неумышленно занесла вирус в изучаемую ими страну третьего мира. В своей области они были экспертами и руководствовались благими намерениями; полагали, что оказывают помощь людям, и на деле никогда всерьез не задумывались о возможности отрицательного появления результатов своей работы.

Занесенная ими вирусная инфекция имела ряд вредных последствий: повысился уровень детской смертности, ухудшилось здоровье и благополучие женщин и детей и, что хуже всего, косвенно была подорвана единственная действенная демократическая политическая сила, что усилило позиции правящего страной традиционного деспота. Когда исследователи столкнулись с разрушительными последствиями проведенной ими работы, они увидели в этом, по меньшей мере, свидетельство крушения их лучших надежд; но, принимая во внимание все вместе взятое, они продолжали считать свою деятельность служащей интересам народа и заявляли, что все дело в неадекватности для данной страны стандартов, в соответствии с которыми измерялись эти так называемые разрушения. Возражая критикам, они говорили, что те пытаются наложить «западные» стандарты на культурную ситуацию, в которой такие стандарты не работают. В этой диковинной самозащите они с симпатией были поддержаны правителями страны, что неудивительно; однако голоса людей, про которых *можно было бы* сказать, используя западные стандарты, что они стали жертвами активности исследователей, мало кем были услышаны — что тоже неудивительно.

Приехавшие в страну исследователи не были биологами, намеревавшимися познакомиться ее людей

с новыми культурами

478

риса, не были они и агрохимиками, испробующими новые пестициды, или докторами, проверяющими здесь вакцины, не получив разрешение на их тестирование в Соединенных Штатах. Они были постмодернистскими критиками науки и прочими мультикультуралистами, которые в ходе своих профессиональных исследований культуры и традиционной «науки» этой страны доказывали, что западная наука является только одним из множества равно валидных нарративов и что ей не следует придавать «привилегированный статус» в сравнении с местными традициями. Традициями, которые другие исследователи — ориентирующиеся на «западные» стандарты биологи, химики, доктора и другие специалисты, — стремятся вытеснить. Введенный ими вирус был не макромолекулой, а меме (реплицирующейся идеей); смысл этой идеи состоит в том, что наука была навязана третьим странам «колониальными» силами и она вовсе не стоит того, чтобы подменять ею практики и верования, приведшие третий мир к его нынешнему состоянию.

Я поведал эту историю с целью драматизации вопроса и внесения беспокойства в то, что мне кажется нынешней ортодоксией среди literary относительно такого рода вещей. Вместе с тем мое беспокойство инспирировано реальными случаями, так сказать, истинными сведениями. События такого рода происходят в Индии и в других местах, и о них сегодня все громче говорят многие авторы<sup>1</sup>.

Придуманная мною маленькая история инспирирована словами Эдварда Уилсона в *Atlantic Monthly*, сказанными им не-

<sup>1</sup> Nanda M. The Epistemic Charity of the Social Constructivist Critics of Science and Why the Third World Should Refuse the Offer // *A House Built on Sand: Exposing Postmodernist Myths about Science*. Ed. by N. Koertge. N.Y., Oxford, 1998. P. 286—311; Afshari Reza. An Essay on Islamic Cultural Relativism in the Discourse of Human Rights // *Human Rights Quarterly* (16), 1994. P. 235—276; Okin S. Is Multiculturalism Bad for Women? // *Boston Review*. Oct./Nov. 1997. P. 25—28; Hoodbhoy P. *Islam and Science: Religious Orthodoxy and the Battle for Rationality*. L. and New Jersey, 1991.

479

сколько месяцев назад: «Ученые, считающиеся ответственными за свои слова, не видят пользы в постмодернизме». В действительности все мы, конечно, ответственны за наши слова. Никто из нас, например, не защищен от клеветы и оскорблений, однако большинство из нас, включая ученых во многих и даже в большинстве областей, как правило, не выступают с заявлениями, могущими, независимо от соображений клеветы или оскорбления, даже косвенным образом нанести вред другим. Подручным измерением данного факта является очевидная смехотворность, с какой мы воспринимаем идею страховки от возможных злоупотреблений со стороны литературных критиков, философов, математиков, историков, космологов и др. Что вообще может сделать математик или литературный критик при выполнении своих профессиональных обязанностей, чтобы вызвать нужду в документе, гарантирующем безопасность от их злонамеренных действий? Можно случайно толкнуть студента или уронить книгу на чью-либо голову, но, кроме таких *outré* побочных эффектов, наша деятельность в прагматическом смысле представляется невинной. Кто-то, возможно, так и думает. В областях с более повышенным и очевидным риском существует давняя традиция быть предельно осторожным и принимать на себя особую ответственность, гарантируя недопущение каких-либо вредных результатов (что четко выражено в клятве Гиппократова). Инженеры, зная, что безопасность тысяч людей может зависеть от качества постройки моста, прибегают к множеству разнообразных приемов, нацеленных на ограничение возможности ошибки, — мосты должны быть крепкими и безопасными. Даже экономисты, часто критикуемые за риск, в который они вовлекают жизни других людей, оказываясь в ситуации, когда от них требуют поддержки тех или иных экономических мер, находящихся на рассмотрении правительства или частных лиц, для безопасности стремятся наложить целительные путы на свои собственные исходные посылки. Как правило, они задаются вопросом — и люди ожидают от них, чтобы они его задавали: «А что, если я ошибаюсь?»

480

Работая в других областях знания, мы редко задаемся таким вопросом, поскольку в нашей студенческой и профессиональной деятельности мы занимаемся темами, которые в соответствии с традицией и здравым смыслом не в состоянии в такой мере повлиять на чьи-либо жизни, чтобы об этом стоило беспокоиться. Если моя тема — удалось ли или не удалось Властосу представить лучшую интерпретацию «Парменида» Платона, или как торговля шерстью повлияла на поэтическое воображение времен Тюдоров, или как понимается время в лучшей версии линейной теории, или как в топологии перевести доказательство в некий новый формализм, — и если в сказанном мною содержится ошибка и даже очень серьезная ошибка, единственный вред будет нанесен только моей собственной академической репутации. Но когда мы хотим оказать большее воздействие на «реальный» мир (в противоположность «академическому»), а многие философы сегодня стремятся к этому, нам нужно принять позиции и правила, действующие в прикладных дисциплинах. Мы должны нести ответственность за наши слова, осознавая, что эти слова, если в них поверят, могут иметь глубокие последствия — как хорошие, так и вредные.

### Истина против преходящих фантазий

Когда я был молодым и неопытным профессором философии, ко мне как-то зашел коллега с факультета сравнительной литературы — известный и модный теоретик литературы — с просьбой о профессиональной помощи. Я был польщен и готов был оказать ему услугу, но ход его вопросов о различных философских сюжетах привел меня в замешательство. В конце концов он решился прямо поведать о причине визита ко мне. Он сказал, что ему нужна «какая-нибудь эпистемология». Буквально так и сказал — «какая-нибудь эпистемология». По-видимому, в этот сезон каждому уважающему себя литературному теоретику приличествовало щеголять какой-нибудь эпистемологией и без нее он чувствовал себя голым, а чтобы

481

одеться, он обратился ко мне с просьбой об эпистемологии, уверенный, что это самый последний писк моды, и ему хотелось бы быть *dernier cri* в эпистемологии. Здравость, обоснованность или (кто-то может сказать) *истинность* такой эпистемологии его не волновали; от нее требовалось только, чтобы она была новой, необычной и модной. Облачайся в нее, иначе тебя не заметят на важном приеме.

В тот момент я осознал, о чем раньше только смутно догадывался, насколько велика пропасть между нами. Поначалу она показалась мне пропастью между серьезностью и фривольностью. Однако поднимавшая во мне волна ощущений собственной правоты все же была наивной реакцией. Мое чувство возмущения, чувство, что этот человек с его эксцентричным проектом отнял у меня время, было столь же немудрым, как и реакция человека, впервые посетившего театр и кинувшегося на сцену в порыве защитить героиню от злодея. «Неужели вы не понимаете? — говорим мы ему поучительно. — Это следует просто *взять на веру*. Это *искусство*. В нем не *предполагается* принимать происходящее на сцене буквально». Если понимать вопросы моего теоретика литературы в таком контексте, в конце концов, они его не дискредитировали. Наверное, я не оскорбился, если бы коллега с факультета драмы обратился с просьбой позаимствовать несколько ярдов моих книг в качестве реквизита мебели в его постановке драмы Тома Стоппарда «Прыгуны» («Jumpers»). Что вообще было бы дурного в экипировке теоретика литературы подходящим набором экстравагантных эпистемологических доктрин, щеголяя которыми он мог бы внести приятное оживление в разговор с коллегами или поразить их?

Плохо было бы то, что, поскольку этот человек не видел наличие пропасти — даже не догадывался о ее существовании, — моя уступчивость его рыночной шалости помогла бы подрыву ценности товара и способствовала эрозии ценного различия. Многие люди, и сторонние наблюдатели и участники, не видят пропасти между серьезностью и фривольностью или активно отрицают ее наличие, и именно в этом суть про-

482

блемы. Печальный факт состоит в том, что в некоторых интеллектуальных кругах, куда входит ряд



наших наиболее продвинутых мыслителей в сфере гуманитаристики и искусства, данная позиция выдается за софистичную оценку тщетности доказательств всех заявок на знание по причине релятивности последнего. На деле за этим далеким от софистичности мнением скрывается высшее проявление наивности, ставшее возможным только благодаря упрямому нежеланию принимать в расчет оправдавшие себя методы научного поиска истины и их силу. Подобно многим другим наивным людям, эти мыслители, рассуждая о демонстрируемой неспособности *их методов* поиска истины в достижении стабильных и ценных результатов и опираясь на свой собственный опыт, наивно обобщают их в том духе, что никому другому вообще неведома возможность открытия истины.

К моему глубокому сожалению, среди людей, вносящих лепту в распространение этой позиции, находится мой друг Дик Рорти. Более четверти века между Ричардом Рорти и мною велась конструктивная полемика. Рорти открыл новые горизонты современной философии, пронизательно показав нам, философам, множество вещей, касающихся истоков наших проектов в философских проектах далекого и близкого времени, в то же время он жестко описывает и предписывает будущие шаги, которые нам следует предпринять. Есть один пункт, относительно которого между нами нет никакого согласия — пока не согласны, — и он связан с многолетним стремлением Рорти показать, что дебаты философов об Истине и Реальности стирают пропасть, а на деле дают лицензию на скатывание в некоторую форму релятивизма. В итоге он сообщает нам, что все эти дебаты не что иное, как «разговоры», а выбор в них той или иной роли определяется только политическими, или историческими, или эстетическими основаниями.

Рорти много раз пытался вовлечь меня в свою компанию, убеждая, что в моих собственных работах можно найти тот или иной взрывоопасный пункт, могущий быть использованным в его проекте разрушения иллюзии достижения объек-

483

тивности. Одним из его цитируемых и любимых пассажей является концовка моей книги «Объясненное сознание» (1991).

«Это только война метафор, вы говорите, но метафоры не являются «только» метафорами; метафоры являются инструментами мысли. Никто не может без них мыслить о сознании, поэтому очень важно вооружить себя лучшим набором из имеющихся в распоряжении инструментов. Посмотрите вокруг, что мы выстроили с помощью этих инструментов. Можете ли вы вообразить все это без них?»<sup>1</sup>

По поводу этого пассажа Рорти говорит: «Мне хочется, чтобы он (Деннет — *Н. Ю.*) сделал еще один шаг вперед и добавил бы к сказанному, что инструменты — это все, что когда-либо может дать исследование, поскольку исследование никогда не является «чистым» в смысле «проекта чистого исследования» Бернарда Уильямса. Оно всегда есть вопрос достижения чего-то, что нам хочется»<sup>2</sup>. Но я никогда не предприму этот шаг; метафоры действительно являются незаменимыми инструментами мысли, однако их нельзя считать единственными. И микроскопы, и математика, и сканеры — тоже такие же инструменты. Конечно, всякое исследование представляет собой приобретение нами чего-то, к чему мы стремимся; но стремимся мы иметь истину о чем-то значимом для нас, если все идет как положено.

Споря об истине, философы спорят о способах недопущения инфляции истины об истине в Истину об Истине, выступают против превращения ее в некую абсолютистскую доктрину, накладывающую непререкаемые требования на наши системы мысли. То же самое можно сказать, например, и относительно дебатов о реальности времени или реальности прошлого. Прделаны глубокие, софистичные и философски ценные исследования темы, является ли, собственно говоря,

<sup>1</sup> Dennett D. *Consciousness Explained*. P. 455.

<sup>2</sup> Rorty R. *Holism, Intrinsicity, and the Ambition of Transcendence* // *Dennett and His Critics. Demystifying Mind*. Ed by Bo Dahlbom. Cambridge Mass., 1993. P. 198.

484

прошлое реальным. Мнения разделились, но вы совершенно превратно поймете суть этих разногласий, предположив, что они подрывают утверждения типа следующих:

жизнь впервые возникла на этой планете более чем три тысячи миллионов лет назад; Холокост имел место во время Второй мировой войны; Джек Руби выстрелил и убил Ли Харви Освальда в 11.25 по далласскому времени 22 ноября 1963 года.

Все это истины о реально происшедших событиях. Их отрицание — ложь. Ни один здравый философ никогда не думал иначе, хотя в пылу споров иногда делались заявления, которые можно было бы интерпретировать таким образом.

Деятельность Ричарда Рорти заслуживает большой и увлеченной его идеями читательской аудитории в сфере искусства, гуманитаристики, в «гуманистических» социальных науках. И все же, когда ревностные читатели интерпретируют его как вдохновителя их постмодернистского скептицизма в отношении истины, они скатываются на путь, идти по которому сам он воздерживается. Когда я его припер к стенке по вопросу истины, он признал, что действительно имеется полезное понятие истины, остающееся непоколебимым и после всех проделанных философами вездыханными опровержений. Рорти согласился считать это подручное скромное понятие истины практически полезным при сравнении, например, двух карт сельской местности на предмет их соответствия местности или при решении вопроса о виновности или невинности лица в совершении преступления.

Даже Ричард Рорти признает пропасть — и ее важность — между явлением и реальностью, между театральными экзерсисами, способными развлекать нас без претензии на сообщение истины, и изысканиями, нацеленными на поиск истины и часто находящими ее. Результат последних изысканий он называет «вегетарианским» понятием истины. Прекрасно, но тогда давайте все мы станем вегетарианцами в отношении истины. Ученые, во всяком случае, никогда не стремились доводить все до крайности.

485

### Базисный инстинкт

Давайте зададимся вопросом об источниках или фундаменте этого скромного, непротиворечивого, вегетарианского понятия истины.

Пока я сейчас говорю, миллиарды организмов на нашей планете вовлечены в игру: спрячься и ищи (hide and seek). Для них это не только игра. Это дело жизни и смерти. В течение более чем трех миллионов лет импульс *«делай все правильно, не допускай ошибки»* был первостепенным для каждого живого организма на этой планете. Поэтому организмы эволюционировали тысячами различных путей, исследуя мир, в котором они живут, различая друзей и врагов, принятие пищи и принятие особи другого пола и игнорируя большую часть всего остального. Для организмов важно иметь адекватную информацию об окружении и ничего более существенного, чем это, у них нет, — но они, как правило, не проводят это через сознание. Они — пользователи благодетелей, получаемых от средств, эксклюзивно детерминированных их дизайном для различения правильного или неправильного для них, однако в случае сбоя инструментария и ложного восприятия ситуации у них, как правило, нет ресурсов для обнаружения сбоя, не говоря уже об исключении ошибки. Они делают свою работу бессознательно. *Распознавание* различия между явлением и реальностью — человеческое открытие. Ряд других видов — некоторые приматы, некоторые щетинистые, может быть, даже некоторые птицы — имеют какие-то признаки осознания феномена «ложного верования» — *делать что-то неправильно*. Они проявляют чувствительность к ошибкам других особей и, может быть, даже некоторую чувствительность к их собственным ошибкам как ошибкам. Однако у них нет способности к рефлексии, необходимой для пребывания в этой возможности, и поэтому они не могут использовать данную чувствительность в форме особого дизайна для ремонта или улучшения их собственных механизмов поиска и механизмов укрытия от опас-

486

ности. Тип соединения пропасти между явлением и реальностью — складочка, которую смастерили только мы, человеческие существа.

Мы относимся к виду существ, открывших сомнение. Достаточно ли пищи заготовлено на зиму? Не просчитался ли я? Подходит ли мне партнер? Надо ли перемещаться на юг? Безопасно ли входить в пещеру? У других существ тоже наблюдается неуверенность относительно таких же вопросов, но поскольку они *не могут задавать самим себе эти вопросы*, не могут самим себе артикулировать свои затруднения, они не в состоянии предпринимать шаги для поиска истины. Они застряли в мире явлений, пытаюсь сделать максимально из того, как им видятся вещи.

Только мы наделены сомнением, и только нас гложет эпистемический зуд поиска лекарства от него — лучших методов приближения к истине. В стремлении найти самый эффективный способ сохранения наших пищевых запасов, наших территорий, наших семей и распознавания путей наших врагов мы обнаружили преимущества разговора об этих вещах с другими, в задавании им вопросов и передаче полезных знаний. Мы изобрели культуру. Затем мы изобрели измерение, арифметику, карты, письмо. Смысл задаваемых вопросов — получить правильные ответы; смысл измерения — *измерять точно*; смысл нарисованных карт — *найти свой путь* к точке назначения. Земля Лжецов существует только в загадках философов: у людей не сложилась традиция системы Ложного Календаря для неверной записи проходящего времени. Короче говоря, нацеленность на истину безоговорочно присутствует в любой человеческой культуре.

Будучи человеческими существами, мы используем коммуникативные навыки не только для сообщения истины, но и для обещаний, угроз, заключения сделок, приглашений, мистификаций, введения в гипнотический транс и просто для надувательства, однако доминантой всех этих действий является сообщение истины, и для этого мы изобрели еще лучшие

487

инструменты. Параллельно с нашими орудиями для сельского хозяйства и транспорта нами создана технология истины — наука. Попробуйте провести прямую линию или нарисовать окружность «от руки». Если у вас нет большого художественного таланта, результат не будет впечатлять. С прямоугольником и компасом вы практически можете элиминировать источники человеческой вариабельности и в то же время получить ясный, объективный результат.

Действительно ли линия является прямой? Насколько она прямая? Пытаясь ответить на эти вопросы, мы разработали еще более чистые тесты, а затем тесты, проверяющие аккуратность этих тестов и т.д., двигаясь ко все большей точности и объективности. Ученые движимы побуждением исходить из базовых мотивов, но они также подвержены мыслить, исходя из желаемого, и бывают столь же корыстными, левоверными и забывчивыми, как и все другие люди. Ученые не причисляют себя к святым; они вовсе не претендуют на роль священников (борясь с человеческими импульсами и моральной неустойчивостью, священники, согласно традиции, делают свою работу лучше, чем все остальные). Ученые считают себя столь же слабыми и подверженными ошибкам, как и все другие люди, но, осознавая наличие источников ошибок в самих себе и в группах, к которым они принадлежат, они разработали системы, связывающие им руки, насильно предохраняющие их результаты от влияния их же слабостей и предрассудков.

Для предотвращения человеческих ошибок созданы системы, не ограничивающиеся только применением физических инструментов. Организация методологических средств тоже находится под влиянием жесткого селективного отбора, подтверждающего их надежность и объективность. Классический пример — двойной слепой эксперимент, в котором во время тестирования и испытуемые, и экспериментаторы лишены возможности знать, кому дается лекарство, а кому плацебо, с тем чтобы ничьи стремления и уловки не влияли на оценку результатов. Статистический результат единичных эксперимен-

488

тов и серии экспериментов воплощается в более широкой практике рутинных повторов результатов индивидуальными исследователями, что в дальнейшем складывается в традицию — хотя и не очень устойчивую — публиковать как позитивные, так и негативные результаты.

Веру в арифметику вдохновляет тот факт, что сотни любителей цифири, решающих одну и ту же проблему, независимо друг от друга получают один и тот же ответ (за исключением тех немногих математиков, которых можно не принимать в расчет, чьи ошибки обнаруживаются и

идентифицируются к взаимному удовлетворению всех). Непревзойденную объективность можно обнаружить и в геометрии, и в других областях математики, служивших со времен античности моделью достоверного знания в отличие от знания мира текучести и противоречий. В «Мено», раннем диалоге Платона, Сократ и мальчик-раб вместе разработали особый случай теоремы Пифагора. Пример Платона свидетельствует об открытом признании стандарта истины, к которому апеллировали все искатели истины, стандарта, который никогда не был подвергнут серьезной атаке и был молчаливо принят, — на деле большинство ревностных оппонентов науки опирались на него даже в вопросах жизни и смерти. (Известна ли вам церковь, управляющая паствой и ее жертвованиями без использования преимуществ арифметики?)

Все это верно, конечно. Вместе с тем наука почти никогда не выглядит непротиворечивой, четкой и сухой, как арифметика. Соперничающие научные группировки часто вовлекаются в пропагандистские битвы, не менее яростные, нежели в политике и в религиозных конфликтах. Порою ярость защитников научной ортодоксии, отстаивающих их доктрины от посягательств еретиков, не имеет аналога в других сферах человеческих риторических битв. Конкуренция за сторонников — и, конечно, за финансирование — нацелена на привлечение к себе внимания и при хорошей организации, как правило, бывает успешной. Побочный ее эффект состоит в том, что война в критических точках науки отвле-

489

кает внимание от наличия огромного неоспоримого фундамента, от самой тяжелой и многотрудной работы, придающей ей силу. В жарких дебатах обычно не говорится о самом важном — о существовании организованного энциклопедического собрания огромного количества согласованных научных фактов.

Роберт Проктор обратил внимание на полезность различения нейтральности и объективности<sup>1</sup>. Геологи, пишет он, знают значительно больше о нефтесодержащих сланцах, чем о других породах, по очевидным экономическим и политическим причинам, однако о нефтеносных сланцах они знают объективно. И многое из знания о нефтеносных сланцах может быть обобщено и перенесено на породы, меньше привлекавшие внимание. Нам хотелось бы, чтобы наука была объективной; но нам не следует стремиться к нейтральности науки. Биологи знают значительно больше о плодовой мушке *дрозофиле* не потому, что она приносит нам богатство, а потому что работая с ней легче получить знание, нежели при работе с другими насекомыми. Их знание о москитах значительно больше, нежели о других насекомых, но в этом случае это связано с большим вредом, наносимым людям москитами, нежели другими насекомыми, возможно, легче поддающимся изучению. Есть множество таких причин для концентрации внимания в науке на каком-нибудь объекте, за которыми скрывается факт, что пути исследования далеки от нейтральности; вместе с тем отсутствие нейтральности в целом не делает эти дороги чем-то менее объективным. Будьте уверены, что временами то одно, то другое девиантное движение ведет к разрушению канонов научной деятельности. Изучение модели болезни у мужчин, например, отвлекаясь от сбора данных об этой же самой болезни у женщин, не только просто нейтрально; это плохая наука, уязвимая в научном и политическом отношениях.

<sup>1</sup> Proktor R. Value — Free Science? Cambridge Mass., 1991.

490

Совершенно верно, что научные ортодоксии давних времен порождали политику, которая, оглядываясь на прошлое, имела серьезные изъяны. Кто-то может симпатизировать, например, позиции Ашиза Нанди, редактора страстной антинаучной антологии «Наука, гегемония и насилие: реквием модернизму»<sup>1</sup>. Пройдя через движения «Атом для мира» («Atoms for Peace») и «Зеленая революция» (упомяну только две из наделавших больше всего шума широких акций, серьезно подорвавших общества третьего мира), Нанди пишет, что «адаптация в Индии устаревших западных технологий рекламируется и подается как великий научный прогресс, даже когда такие адаптации превращают все дисциплины или сферы знания в простые интеллектуальные машины для адаптации, воспроизведения и тестирования уже не имеющих спроса на Западе моделей, которые там часто воспринимаются как слишком опасные или экологически нежизнеспособные»

(с. 8). И все же нам следует оценивать такие явления как политические злоупотребления, а не как фундаментальный изъян самой науки.

### Как доминирует наука

Методы науки не застрахованы от глупости, тем не менее они, безусловно, предпочтительнее. Самое важное состоит в следующем: существует традиция критики, заставляющая при обнаружении изъянов везде и всегда заниматься усовершенствованием. *Методы* науки, как и все прочее под Луной, сами по себе являются объектами научной дотошности, в результате чего метод становится методологией, анализом методов. В свою очередь, *методология* попадает в сферу внимания *эпистемологии*, исследования самого исследования, и вне сферы научного вопрошания не остается ничего. Ирония состоит в том, что плоды научной рефлексии, свидетельству-

<sup>1</sup> Nandy A. *Science, Hegemony and Violence: A Requiem for Modernity*. Delhi, Oxford, 1988.

491

ющие о неэлиминируемых темных пятнах несовершенства, время от времени используются теми, кто подозрителен к науке как их основе, с целью отрицания ее привилегированного статуса в департаменте поиска истины — как будто институты и практики, выдвигаемые в качестве конкурентов науки, являются в этом отношении лучше. Но где вы найдете образцы религиозной ортодоксии, легко отброшенной перед лицом неопровержимой очевидности? В науке же вчерашние ереси снова и снова становятся сегодняшними новыми ортодоксиями. В истории ни одной религии вы не найдете примера такой модели.





## Приложение 2. Библиография

### Книги Д.К. Деннета

Content and Consciousness. Routledge & Kegan Paul, N.Y., 1969, xii, 198 p. (2<sup>nd</sup> ed. L, Boston, 1986).  
 Brainstorms. Philosophical Essays on Mind and Psychology. Montgomery Vt: Bradford Books, 1978, XXII, 353 p. (2<sup>nd</sup> and 3<sup>rd</sup> ed. 1986, 1999):  
 The Mind's I: Fantasies and Reflections on Self and Soul (Compauned and ed. with Hofstadter D.R.), Basic Books, N.Y., 1981, p. VII, 501 p.  
 Elbow Room: the Varieties of Free Will Worth Wanting. Oxford. Crarendon Press, N.Y., 1984, 200 p.  
 The Intentional Stance. Cambridge MA, Bradford Book, L., 1989, XI, 388 p.  
 Consciousness Explained. Boston, Littl Brown and Co, 1991 and L.: Allen Lane, XIII, 511 p.  
 Darwin's Dangerous Idea. Evolution and the Meanings of Life. N.Y., Simon and Schuster, L., 1995, 586 p.  
 Kinds of Minds. Toward an Understanding of Consciousness. Basic Books, N.Y., 1996, VII, 184 p.  
 Brainchildren. Essays on Designing Minds. Cambridge MA, Bradford Book, 1998, IX, 418 p.  
 Freedom Evolves. Viking, N.Y., Penguin Books, L., 2003, XIII, 347 p.

493

### Тексты Деннета на русском языке

**Деннет Д.К.** Виды психики: на пути к пониманию сознания. Под. ред. Л.Б. Макеевой. Пер. с англ. А. Веретенникова. М., 2004;  
 Онтологическая проблема сознания. Пер. с англ. А.Л. Блинова // Аналитическая философия: становление и развитие. М., 1998. С. 361—375;  
 — Условия личностного. Пер. с англ. Н.С. Юлиной // История философии (ежегодник), М., 1999. № 5. С. 199—223;  
 — Почему нам важно понимать это правильно. Постмодернизм и истина. Пер. с англ. Н.С. Юлиной // Вопросы философии, М., 2001. № 8. С. 93—100;  
 — Почему каждый из нас является новеллистом. Пер. с англ. Н.С. Юлиной // Вопросы философии, 2003. № 2. С. 121—130.

### Список статей, рецензий, интервью Деннета\* (Работы с 1987 по 2003 гг. доступны в Интернете).

Адрес: <http://ase.tufts.edu/cogstud/pubpage.htm> ). Работы с 1968 по 1986 гг. имеются только в печатном виде.

2003

**Dennet D.C.** Has Darwin's Theory of Natural Selection Transformed our View of Humanity's Place in the Universe? // Life:

\* А) Поскольку статьи, рецензии, выступления Деннет часто помещает на сайты в Интернете не только в виде опубликованных, но и готовящихся к публикации черновых набросков (drafts), мы даем их в том библиографическом виде, в каком они фигурируют там, то есть без принятого у нас оформления.

Б) Представленный ниже список публикаций Деннета может служить хорошей иллюстрацией к нашему наблюдению, что характерными чертами его философии (и современной англоязычной философии в целом) является дискуссионность, нескончаемый полилог с оппонентами и коллегами.

494

The Science of Biology, Ed. by D.E. Purves, G.H. Sadava, and H.C. Heller, 7th Edition, Sinauer Associates, Inc Publishers;

— «Shame on Rea». A reply to Michael C Rea's «Dennett's Bright Idea»;

- On Failures of Freedom & the Fear of Science // Daedalus, Journal of the American Academy of Arts, Winter 2003. P. 126—130;
- Making Ourselves at Home in Our Machines. Review of Daniel Wegner: «The Illusion of Conscious Will». Cambridge MA, 2002;
- The Mythical Threat of Genetic Determinism // The Chronicle of Higher Education, January 31, 2003;
- What RoboMary Knows // Final draft for: «Knowledge Argument». Ed. by Torin Alter;
- Who's On First? Heterophenomenology Explained // Journal of Consciousness Studies, 10, № 9—10, 2003. P. 10—30;

## 2002

- Forestalling a Food Fight over Color. Commentary on Byrne & Hilbert, «Color Realism and Color Science» // Behavioral and Brain Sciences;
- Could There be a Darwinian Account of Human Creativity? Final draft for «Evolution, From Molecules to Ecosystems», Ed by. A. Moya and E. Font, Oxford;
- Look out for the Dirty Baby. Review on B. Baars, «The Double Life of B.F.Skinner...», June 10, 2002 // Journal of Consciousness Studies, forthcoming;
- Does Your Brain Use the Images in It, and if so, How? Commentary on Pylyshyn // Behavioural and Brain Sciences;
- How Could I be Wrong? How Wrong Could I Be? // Journal of Consciousness Studies (Special issue) on «The Grand Illusion», January 13, 2002;
- The Baldwin Effect: a Crane, not a Skyhook // Evolution and Learning: The Baldwin Effect Reconsidered. Ed. by B. Weber and D. Depew, Cambridge MA, 2002, in press;
- Higher Order Truths About Chmess;

495

## 2001

- The Fantasy of First-Person Science. (A written version of a debate with David Chalmers, held at Northwestern University, Evanston, IL, February 15, 2001, supplemented by an E-mail debate with Alvin Goldman);
- Consciousness: How Much is that in Real Money? // R. Gregory, ed., Oxford Companion to the Mind, on consciousness, Dec. 12, 2001;
- Altruists, Chumps, and Inconstant Pluralists. Commentary on Sober and Wilson, Unto Others: The Evolution and Psychology of Unselfish Behavior, Sept. 22, 2001;
- «Surprise, surprise». Commentary on O'Regan and Noe // Behavioural and Brain Sciences, 2001. V. 24. № 5. P. 982;

## 2000

- From Typo to Thinko... // Fyssen volume on cultural evolution;
- and Christopher Taylor. Who's Afraid of Determinism...? // Handbook of Free Will. Ed. by R. Kane, CUP;
- «Foreword» // R. Aunger «Darwinizing Culture», 2000;
- In Darwin's Wake, Where am I? // APA Presidential Address, Dec. 29, 2000;
- Review of E. Avital and E. Jablonka, «Animal Traditions: Behavioural Inheritance in Evolution». Cambridge Un. Press, 2000 // Journal of Evolutionary Biology;
- Why Getting It Right Matters. Postmodernism and Truth // Free Inquiry. V. 20. № 1. 1999/2000. P. 40—43;
- Are we Explaining Consciousness Yet? // «Cognition». August 27, 2000;
- Self Portrait // Companion to the Philosophy of Mind. Ed. by Guttenplan, Oxford;
- The Case for Rorts // Rorty and His Critics. Ed. by R.B.Brandom, Oxford, 2002. P. 91—101;

496

## 1999

- with Viger Ch. Is Hirsch or Wilson confused? A commentary on: «The Pitfalls of Heritability»;
- The Possibility of Subisomorphic Experiential Differences. Commentary on Palmer // Behavioral and

Brain Sciences, by Ch. Viger;

— Swift and Enormous. Commentary on Palmer // Behavioural and Brain Sciences;

— and Densmore Sh. The Virtues of Virtual Machines // Philosophy and Phenomenological Research, Sept., 1999. V. LIX. № 3;

— It's Not a Bug, It's a Feature // Journal of Consciousness Studies. V. 7. № 4. 2000. P. 25—27;

— and Viger Ch. Sort-of Symbols? Commentary on Barsalou «Perceptual Symbol Systems» // Behavioural and Brain Sciences. V. 22. № 4. Aug. 1999. P. 613;

— We Earth Neurons. August 15, 1999;

— The Zombie Hunch: Extinction of an Intuition? // Final draft for the Royal Institute of Philosophy, Millennial Lecture. Oct. 1999;

— «Foreword» for Nicholas Humphrey, «Leaps of Faith: Science, Miracles, and the Search for Supernatural Consolation». N.Y., 1999;

### 1998

— and Westbury C.F. Mining the Past to Construct the Future: Memory and Belief as Forms of Knowledge // Memory, Brain, and Belief. Ed. by D. Schacter, and E. Scarry, Cambridge MA (in press);

— Making Tools for Thinking // Metarepresentation. Ed. by D. Sperber (volume on conference at Simon Fraser University), Feb. 7—8, 1997; Final draft. April 10, 1998;

— Things about Things. Final draft for Lisbon Conference on Cognitive Science. May 1998;

— and C.F. Westbury. Stability is Not Intrinsic. Commentary on O'Brien and Opie, «A Connectionist Theory of Phenomenal Experience» // Behavioral and Brain Sciences. 1998;

497

— Postmodernism and Truth // Final draft for World Congress of Philosophy. August 13, 1998;

— Memes: Myths, Misunderstandings and Misgivings // Draft for Chapel Hill Lecture. Oct. 1998;

### 1997

— Faith in the Truth. Final draft of Amnesty Lecture. Oxford, February 18, 1997;

— The Evolution of Evaluators // For International School of Economic Research, volume on evolutionary economics, University of Siena. Summer, 1997;

— Reply to Mulhauser // Philosophical Books. V. 38. P. 89—92. April 1997;

— Snowmobiles, Horses, Rats, and Memes. A comment on «A Critique of Evolutionary Archeology», by J.L. Boone and E.A. Smith // Current Anthropology, forthcoming;

— The Case of the Tell-Tale Traces: A Mystery Solved; a Skyhook Grounded. Comment on Behe // For Notre Dame conference. April 5, 1997;

— Appraising Grace: What Evolutionary Good is God? Review of W. Burkert, «Creation of the Sacred: Tracks of Biology in Early Religion» // The Sciences. Jan/Feb, 1997;

### 1996

— Review of «The Ancient Mind: Elements of Cognitive Archeology», ed. by C. Renfrew and E.B.W. Zubrow. Cambridge MA, 1994;

— Granny versus Mother Nature — No Contest. Response to Fodor (on DDI);

— and Densmore Sh. The Virtues of Virtual Machines. Dec. 20, 1996;

— Consciousness: More Like Fame than Television // Die Technik auf dem Weg zur Seele. Munich, 1996;

498

— Quantum Incoherence. Review of A.G. Cairns-Smith «Evolving the Mind: on the Nature of Matter and the Origin of Consciousness» // Nature. V. 381. 6 June, 1996;

— Facing Backwards on the Problem of Consciousness. Commentary on D. Chalmers // Journal of Consciousness Studies. V. 3. № 1 (special issue, part 2). 1996. P. 4—6;

— Producing Future by Telling Stories // From: The Robot's Dilemma Revisited,.. Eds. Ford, Pylyshyn, Ablex;

— Review of Hofstadter & F.A.R.G «Fluid Concepts And Creative Analogies» // Complexity Journal, Santa Fe Institute. V. 1. № 6. 1995/96. P. 9—12;

- Review of Thomas Nagel's «Other Minds: Critical Essays, 1969—1994», OUP, 1995 // *Journal of Philosophy*. V. XCIII. № 8. Aug. 1996. P. 425-428;
- The Myth of Double Transduction. Final Draft // For TUCSON II volume, Stuart Hameroff, ed. Sept. 30, 1996;
- Reflections on Language and Mind // For Sheffield Conference volume. October 7, 1996;

### 1995

- Is Perception the «Leading Edge» of Memory? // A. Spadafora (ed.) *Iride: Luoghi della memoria e dell'oblio*. anno. VIII, № 14. April. P. 59—78;
- Review of Antonio Damasio, *Descartes' Error* // *Times Literary Supplement*. August 25. P. 3—4;
- Animal Consciousness: What Matters and Why // *Social Research*. V. 62. № 3. 1995. P. 691—710;
- Evolution as an Algorithm — the Ultimate Insult // *The Mind, the Brain and Complex Adaptive Systems*. Ed. by H. Morowitz and J. Singer. Santa Fe Inst. *Studies in the Sciences of Complexity, Proceedings*. V. XXII. Addison-Wesley. 1995. P. 221—223;
- Seeing is Believing — or is it? // *Perception*. Ed. by Akins, *Vancouver Studies in Cognitive Science*. V. 5. Oxford. P. 158—172;

499

- Overworking the Hippocampus. Commentary on Jeffrey Gray // *Behavioral and Brain Sciences*. V. 18. № 4. 1995. P. 677—678;
- The Path not Taken. Commentary on: Ned Block, «On Confusion About a Function of Consciousness» // *Behavioral and Brain Sciences*. V. 18. № 2. 1995. P. 252—253;
- Do Animals have Beliefs? // Herbert Roitblat & Jean-Arcady Meyer, eds., *Comparative Approaches to Cognitive Sciences*. Cambridge MA, 1995. P. 111—118;
- How to Make Mistakes // *How Things Are*. Ed. by J. Brockman and K. Matson. N.Y., 1995. P. 137—144;
- The Unimagined Preposterousness of Zombies. Commentary on T. Moody, O. Flanagan and T. Polger // *Journal of Consciousness Studies*. V. 2. № 4. 1995. P. 322—326;

### 1994

- Instead of Qualia // *Consciousness in Philosophy and Neuroscience*. Ed by Revonsuo A. and Kamppinen M. Hillsdale. NJ., 1994. P. 129—139;
- Dennett and Carr Further Explained: an exchange // Emory Cognition Project, Report 28, Department of Psychology, Emory University. Apr;
- Artificial Life as Philosophy // *Artificial Life*. V. 1. № 1. 1994;
- Cognitive Science as Reverse Engineering: Several Meanings of «Top Down» and «Bottom Up» // *Logic, Methodology and Philosophy of Science*, IX, D. Prawitz, B. Skyrms and D. Westersteel, eds., Elsevier Science, BV. Amsterdam, North-Holland. P. 679—689;
- Self Portrait // *A Companion to the Philosophy of Mind*. Ed. by S. Guttenplan, Blackwell Press. Oxford. P. 236—244;
- «Real Consciousness» and «Instead of Qualia» // *Consciousness in Philosophy and Cognitive Neuroscience*. Ed. by A. Revonsuo & M. Kamppinen, Hillsdale, N.J.: Lawrence Erlbaum;

500

- and M. Kinsbourne. Counting Consciousnesses: None, One, Two, or None of the Above? Commentary on «Time and the Observer» // *Behavioral and Brain Sciences*. V. 17. № 1. P. 178—180;
- The Practical Requirements for Making a Conscious Robot // *Philosophical Transactions of the Royal Society*, A. V. 349. P. 133—146;
- Review of Jackendoff's «Patterns in the Mind» and Pinker's «The Language Instinct» // *London Review of Books*. 23 June. V. 16. № 12. P. 10—11;
- The Role of Language in Intelligence // *What is Intelligence? The Darwin College Lectures*. Ed. J. Khalfa, Cambridge;
- Interview with Daniel Dennett // *Journal of Cognitive Neuroscience*. V. 7. № 3;
- Get Real. Reply to my critics // *Philosophical Topics*. V. 22. № 1&2. Spring & Fall, 1994;
- E Pluribus Unum? Commentary on Wilson & Sober: Group Selection, BBS. V. 17. № 4. 1994. P.



617—618;

— Consciousness in Human and Robot Minds // Cognition, Computation and Consciousness. Ed. by Ito, et al., OUP, Kyoto. 1994;

### 1993

— Review on Searle J.R. «The Rediscovery of Consciousness» (Cambridge MA, 1992) // Journal of Philosophy. V. XC. № 4. April 1993. P. 193—205;

— Living on the Edge. Symposium on Daniel Dennett's «Consciousness Explained» // Inquiry. V. 36. № 1/2. 1993. P. 135—160;

— Confusion Over Evolution: An Exchange (with S J. Gould) // The New York Review of Books. V. XL. № 1&2. P. 43—44;

— Review of Varela et al, «The Embodied Mind» // American Journal of Psychiatry. V. 106. № 1. Spring. P. 121—125;

501

— Back from the Drawing Board: Reply to my Critics // Dennett and His Critics: Demystifying Mind. Ed. by Dahlbom. Oxford, 1993;

— The Message is: There is no Medium. Reply to Jackson, Rosenthal, Shoemaker & Tye // Philosophy & Phenomenological Research (Dec). V. 53. № 4. P. 889—931;

— Review of Allen Newell «Unified Theories of Cognition» // Artificial Intelligence. V. 59. № 1—2. P. 285—294;

— Down with School! Up with Logoland! Review of Seymour Papert, «The Children's Machine: Rethinking School in the Age of the Computer» // New Scientist, 140, 1988. Nov. 1993. P. 45—46;

— and M. Kinsbourne. Multiple Drafts: An Eternal Golden Braid. Response to Glicksohn and Salter // Behavioral and Brain Sciences. V. 18. № 4. 1995. P. 810—811;

— Learning and Labeling. Commentary on Clark and Karmiloff-Smith «The Cognizer's Innards» // Mind and Language. V. 8. № 4. P. 540—547;

— Evolution, Teleology, Intentionality (continuing commentary on «Précis of The Intentional Stance») // Behavioral and Brain Sciences. V. 16. № 2. P. 289—391;

— Caveat Emptor. Reply to Essays on «Consciousness Explained» (reply to Mangan, Toribio, Baars and McGovern) // Consciousness and Cognition. V. 2. P. 48—57;

### 1992

— The Self as a Center of Narrative Gravity // Self and Consciousness: Multiple Perspectives. Kessel, Cole, Johnson Eds., Erbaum, 1992;

— and Kinsbourne M. Time and Observer: The Where and When of Consciousness in the Brain // The Behavioral and Brain Science. V. 15. 1992. P. 183—241;

— Verbal Language. AICA conference, Oct. 8, 1992, Milan;

— and Kinsbourne, M. Time and the Observer // BBS, 15. P. 183—247; & reply, «Escape from the Cartesian Theater»;

502

— Temporal Anomalies of Consciousness: Implications for the Uncentered Brain // Neurophilosophy and Alzheimer's Disease, Christen et al. eds., Springer-Verlag Berlin Heidelberg;

— Filling In versus Finding Out: A Ubiquitous Confusion in Cognitive Science // Cognition, Conceptual and Methodological Issue. Ed. by Pick, Van den Broek, Knill, American Psychological Association;

— Review of Varela et al, «The Embodied Mind» and Edelman, «Bright Air, Brilliant Fire» // New Scientist, 13 June;

— Hitting the Nail on the Head. Commentary on Thompson, Palacios and Varela // Behavioral and Brain Sciences. V. 15. № 1. 1992. P. 35;

— La Compréhension Artisanale. (French translation of «Do-It-Yourself Understanding») // D. Fiset ed., Daniel C. Dennett et les Strategies Intentionnelles, Montreal, 1992, Lektou, 11, Winter;

### 1991

— Granny's Campaign for Safe Science // B. Lower and G. Rey eds., Meaning in Mind: Fodor and his

Critics, Blackwell;

— Two Contrasts: Folk Craft versus Folk Science and Belief versus Opinion // J.D. Greenwood, ed., *The Future of Folk Psychology*;

— The Brain and its Boundaries. Review of C McGinn, «The Problem of Consciousness» // *The Times Literary Supplement*. May 10, 1991;

— Real Patterns // *Journal of Philosophy*. V. 88. № 1. P. 27—51;

— Lovely and Suspect Qualities // E. Villanueva, ed., *Consciousness*;

### 1990

— Mother Nature versus the Walking Encyclopedia // *Philosophy and Connectionist Theory*. Ed. by Ramsey, Rumelhart and Stich. 1990;

503

— Memes and the Exploitation of Imagination // *The Journal of Aesthetics and Art Criticism*. V. 48. 1990. P. 127—135;

— Teaching an Old Dog New Tricks. Commentary on Schull // *BBS*. V. 13. № 1. March 1990;

— Betting Your Life on an Algorithm. Commentary on Penrose // *BBS*. V. 13. №4. December;

— Review of «Artificial Life», ed. by Langton // *Biology and Philosophy*. V. 5. P. 489-492;

### 1989

— and N. Humphrey. *Speaking for Our Selves* // Raritan;

— The Origins of Selves // *Cogito*. V. 3. Autumn 1989. P. 163—173;

— Review of Penrose «The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics» // *Times Literary Supplement*;

— Cognitive Ethology: Hunting for Bargains or a Wild Goose Chase? // *Goals, Oun Goals and No-Goals*. Ed. by Montefiore and Noble. Unwyn Hyman, 1989;

— Memes and the Exploitation of Imagination // *The David and Miriam Mandel Lecture* // *American Society for Aesthetics*. Oct. 27, 1989. P. 1;

### 1988

— Why Everyone is a Novelist // *Times Literary Supplement*. September 16—22, 1988. P. 96—98;

— Review of J. Fodor «Psychosemantics» // *Journal of Philosophy*;

— When Philosophers Encounter AI // *Daedalus*, 117. P. 283—296;

— The Intentional Stance in Theory and Practice // Whiten and Byrne, eds., *Machiavellian Intelligence*, Oxford University Press;

— Quining Qualia // *Consciousness in Modern Science*. Ed. by A. Marcel and E. Bisiach. Oxford, 1988. P. 42—47;

504

— Why Creative Intelligence is Hard to Find. Commentary on Whiten and Byrne // *BBS*;

— Out of the Armchair and Into the Field // *Poetics Today*;

— Review of Honderich, «A Theory of Determinism: The Mind, Neuroscience, and Life-hopes», «Coming to Terms with the Determined» // *Times Literary Supplement*. November, 4—10;

— Ways of Establishing Harmony // B. McLaughlin, ed., *Dretske and His Critics*;

— Evolution, Error and Intentionality // Y. Wilks and D. Partridge, eds, *Sourcebook on the Foundations of Artificial Intelligence*. New Mexico University Press, 1988;

### 1987

— Eliminate the Middletoad! Commentary on Ewert // *Behavioral and Brain Sciences*. V. 10. 1987. P. 372—374.

— Commentary on Cam // *Philosophy and Phenomenological Research*. V. 48, 1987. P. 339—341;

— The Logical Geography of Computational Approaches: a View from the East Pole // *Problems in the Representation of Knowledge*. Ed. by M. Harnish and M. Brand (eds). University of Arizona Press, 1987;

— Skinner Placed. (Response to U.T. Place, *Skinner Re-Skinned*) // S. and C Modgil (eds). «B.F. Skinner: Consensus and Controversy», Brighton, 1987;

— with Haugeland J. *Consciousness* // *The Oxford Companion to the Mind*. Ed. by R.L. Gregory (ed.),

Oxford University Press, 1987;

— with Haugeland J. *Intentionality // The Oxford Companion to the Mind*. Ed by R.L. Gregory. Oxford University Press, 1987;

— Comments. (On Pylyshyn, Computers, Knowledge, and the Human Mind) // *Creativity and Liberal Learning*. Ed by D. Tuerck, Norwood, N.J., 1987;

505

### 1986

— *Computer Models and the Mind — a View from the East Pole // The Times Literary Supplement*, December 14, 1984. P. 1453—1454;

— *Thinking about Thinking: the Mind as Virtual Machine // The Human Mind — the Human Brain and Beyond*, American Broadcasting Companies, Inc., 1984;

— Review of V. Braitenberg, *Vehicles: Experiments in Synthetic Psychology // Philosophical Review*. V. 95. 1986. P. 137—139;

— Julian Jaynes's *Software Archeology // Canadian Psychology*. V. 27. 1986. P. 149—154;

— *The Imagination Extenders // Humanizing the Computer: a Cure for the Deadly Embrace*, Ed. by D. Flaherty, Belmont, Calif.: Wadsworth, 1986;

— *Engeneering's Baby (commentary of Sayre) // Behavioral and Brain Sciences*. V. 9. 1986. P. 141—142;

— with K. Akins. *Who May I say is Calling? (Commentary on Hoffman) // Behavioral and Brain Sciences*. V. 9. 1986. P. 517—518;

— *Philosophy as Mathematics or as Anthropology // Mind and Language*. V. 1. Spring 1986. P. 18—19;

— *Is there an Autonomous «Knowledge Level»? (Commentary on Newell) // Meaning and Cognitive Structure: Issues in the Computational Theory of Mind*. Ed by Z. Pylyshyn and W. Demopoulos, Norwood, N.J.: Ablex 1986;

— *Information, Technology and the Virtues of Ignorance // Daedalus, Proceedings of the American Academy of Arts and Sciences*. V. 115. Summer 1986. P. 135—153;

### 1985

— *Music of the Hemispheres (Review of M.S. Gazzaniga «The Social Brain») // The New York Times Book Review*. November 17. V. 53, 1985;

506

— *Can Machine Think? // How We Know*. Ed. by M. Shafto. San Francisco, 1985.

— *When Does the Intentional Stance Work? (reply to continuing commentary) // Behavioral and Brain Sciences*. V. 8. 1985. P. 758—766;

— *Why Believe in Belief? (Review of S. Stich, «From Folk Psychology to Cognitive Science: the Case against Belief») // Contemporary Psychology*. V. 30. 1985. P. 949;

— *Commentary in D.G. Bobrow and P.L. Hayes (eds), «Artificial Intelligence — Where Are We?» // Artificial Intelligence*. V. 25. 1985. P. 375—415 (409-410);

### 1984

— *Cognitive Wheels: the Frame Problem of AI // Minds, Machines and Evolution*. Cambridge etc. 1984. P. 153—174;

— *Formulating Human Purposes: Meta-Engineering Computers for People // «Information Processing»*. V. 83. Ed. by R.E.A. Mason. Amsterdam, 1984;

— *Cognitive Weels: the Frame Problem of AI // Minds, Machines and Evolution*. Ed by C. Hookway. Cambridge, 1984;

— *Wishful Thinking (Commentary on Skinner) // Behavioral and brain Science*. V. 7. 1984. P. 556—557;

— «Forword» for R.G. Millican's «*Language, Thought and Other Biological Categories*». Cambridge MA, 1984;

— *Carving the Mind at its Joints (Review of J. Fodor «The Modularity of Mind») // Contemporary Psychology*. V. 29. 1984. P. 285—286;

— *Correspondent's Report: Recent Work in Philosophy AI // Artificial Intelligense*. V. 22. 1984. P. 231—

234.

The Role of Computer Metaphor in Understanding the Mind // Computer Culture, Annals of NY Academy of Science. V. 426. 1984. P. 266—275;

507

— I could Not Done Otherwise — So What? // Journal of Philosophy. V. 81. Oct. 1984. P. 553—565;

1983

— Artificial Intelligence and the Strategies of Psychological Investigation // States of Mind, ed by J. Miller // BBC, L., 1983;

— Intentional Systems in Cognitive Ethology: The «Panglossian Paradigm» Defended // The Behavioral and Brain Science. V. VI. 1983. P. 343—390;

1982

— How to Study Human Consciousness Empirically or Nothing Comes to Mind // Synthese. V. 53. № 2. November 1982. P. 159—180;

— «Beyond Belief» // Thought and Object. Ed. by Woodfield A. Oxford, 1982;

— The Myth of the Computer: An Exchange (Reply to John Searle's Review of «The Mind's I») // New York Review of Books. V. 29, June 24, 1982. P. 56—57;

— Philosophy According to Nozick // New Boston Review. V. 7. 1982. P. 9—10;

— Grey Matter and Mind // Radio Times. 8—14 May, 1982. P. 70—72;

— The Well Furnished Mind (Review of R. Gregory «Mind in Society» // Contemporary Psychology. V. 27. 1982. P. 597—598;

— Why We Think What We Do about Why We Think What We Do: Discussion on Goodman's «On Thought without Words» // Cognition. V. 12. 1982. P. 219—227;

— Comments on Rorty // Synthese. V. 53. № 2. November 1982. P. 349—356;

— The Imagination Extender // Psychology Today. V. 16. 1982. P. 32—39;

— Styles of Mental Representation // Proceedings of the Aristotelean Society. New Series. V. 83. 1982/1983. P. 213—226;

508

1981

— Wandering Where the Yellow Went. Commentary on Sellars's Cams Lecture // The Monist. V. 64. 1981. P. 102—108;

— Three Kinds of Intentional Psychology // Reduction, Time and Reality. Ed. by R.Healey. Cambridge, 1981. P. 37—61;

— True Believers: the Intentional Strategy and Why it Works // Scientific Explanation, ed A.F. Heath. L., 1981; and in Thought and Object: Essays on Intentionality, ed. by A.Woodfield. L., 1981;

— Verbal Language as a communicative system;

— The Kinds of Intentional Psychology // Reduction, Time and Reality. Ed. by Healy R. Cambridge, 1981;

— True Believers: The Intentional Strategy and Why it Works // Scientific Explanation. Ed. by A. Heath, Oxf. Un Press, 1981. P. 53—75;

— Making Sense of Ourselves // Philosophical Topics, V. XII. № 1. 1981. P. 63—81;

1980

— Passing the Buck to Biology (Commentary on Chomsky) // Behavioral and Brain Sciences. V. 3. № 19. 1980;

— The Milk of Human Intentionality // Behavioral and Brain Sciences. V 3. 1980. P. 429;

1979

— Review of Popper and Eccles «The Self and its Brain» // Journal of Philosophy. V. 76. 1979. P. 91—97;

— On the Absence of Phenomenology // Body, Mind and Method. Ed. by D. Gustafson and B. Tapscott. Dordrecht, 1979;

— Review of R. Aquila «Intentionality: a Study of Mental Acts» and E. Casey «Imagining: a

Phenomenological Analysis» // Southwestern Journal of Philosophy. V 9. 1979. P. 139—143;  
 — Artificial Intelligence as Philosophy and as Psychology // Philosophical Perspectives on Artificial Intelligence, Atlantic Highlands, 1979;

509

### 1978

— and Humphrey N. Speaking for Ourselves // Raritan;  
 — Current Issues in the Philosophy of Mind // American Philosophical Quarterly. Pittsburgh, 1978. V. 15. № 4. P. 249—261;  
 — Toward a Cognitive Theory of Consciousness // Perceptions and Cognition. Minnesota Studies in the Philosophy of Science. V. IX. 1978;  
 — Why You Can't make Computer that Feels Pain // Synthese. V. 38. 1978. P. 415—456;

### 1977

— Critical Notice of J. Fodor «The Language of Thought» // Mind. V. 86. 1977. P. 265—280;  
 — Review of W. Matson «Sentience» // International Studies in Philosophy. V. 9. 1977. P. 182—183;  
 — Review of J. Glover (ed.) The Philosophy of Mind // Teaching Philosophy. V. 2. 1977. P. 196—197;

### 1976

— Are Dreams Experiences? // Philosophical Review. V. 85. 1976. P. 151—171;  
 — Conditions of Personhood // The Identity of Persons. Ed. by A. Rorty. Un of California Press, 1976;

### 1975

— Why the Law of Effect Will Not Go Away // Journal for the Theory of Social Behavior. V. 5. 1975. P. 169—187;  
 — Brain Writing and Mind Reading // Language, Mind and Knowledge. Ed by K. Gunderson. Minnesots Studies in the Philosophy of Science. V. VII. 1975;  
 — Review of: G. Warnock and B.F. Skinner. «Behavior Control: Freedom and Morality» // Teaching Philosophy. V. 1. 1975. P. 175—177;

510

### 1974

— Comment on Wilfrid Sellars // Synthese. V. 27. 1974. P. 239—244;

### 1973

— Mechanism and Responsibility // Essays on Freedom and Action. Ed. by T. Honderich. L., 1973;

### 1972

— Review of J.R. Lucas «Freedom of the Will» // Journal of Philosophy. V. 69. 1972. P. 527—531;

### 1971

— Intentional Systems // Journal of Philosophy. V. 68. 1971. P. 87—106;  
 — Review of CO. Ewans, «The Subject of Consciousness» // Philosophical Quarterly. V. 21. 1971. P. 180—181;

### 1968

— Machine Traces and Protocol Statements // Behavioral Science. V. 13. 1968. P. 155—161.  
 — Geach on Intentional Identity // Journal of Philosophy. V. 65. 1968. P. 335—341.  
 — Future of Intentional Action // Philosophy and Phenomenological Research. V. 29. 1968. P. 232—244.



## Литература о философии Деннета на русском языке

Юлина Н.С. Дэниел Деннет: концепция сознания и личностного // История философии (ежегодник). № 5. М., 1999. С. 192—198.

— Деннет versus Рорти // Системные исследования. Методологические проблемы. М., 2001. С. 100—122;

— Деннет о вирусе постмодернизма. Polemica с Р. Рорти о сознании и реализме // Вопросы философии. М., 2001. № 8. С. 78—92.

511

— Дэниел Деннет // Новая философская энциклопедия. М., 2000. С. 623—624;

— Д. Деннет о проблеме ответственности в свете механицистского объяснения человека // История философии. М.: ИФРАН, 2002. № 8. С. 58—77;

— Д. Деннет: самость как «центр нарративной гравитации» или почему возможны самостные компьютеры // Вопросы философии, 2003. № 2. С. 104—121.

## Литература о философии Деннета на английском языке

### Отдельные издания

Daniel C. Dennett et les Strategies Intentionells. Ed. by D. Fiset. Montreal, 1992.

Dennett and His Critics: Demistifying Mind. Ed. by Bo Dahlbom. Oxford, Cambridge MA, 1993.

Daniel Dennett's «Consciousness Explained». Symposium // Inquiry. V. 36. № 1/2. 1993.

Dennett's Philosophy: A comprehensive Assessment. Ed. by D. Ross, A. Brook, D. Thomson. Cambridge MA, 2000.

Daniel Dennett. Ed. by A. Brook and D. Ross. Cambridge, 2002.

Philosophical Topics. V. 22. № 1&2.

### Статьи

Block N. On Daniel Dennett's «Consciousness Explained» // Journal of Philosophy. V. XC. № 4. April 1993. P. 193—205.

Bricke J. Consciousness and Dennett's Intentionalist Net // Philosophical Studies. V. 48. № 2. Sept. 1985. P. 249—256.

Brown A. The Darwin Wars. N.Y., 1999.

Cam Ph. Dennett on Intelligent Storage // Philosophy and Phenomenological Research. V. 45. № 2. Dec. 1984.

Chalmers D. Facing up to the Problem of Consciousness // Journal of Consciousness Studies. V. 2. P. 200—219;

512

— First-person Methods in the Science of Consciousness // Consciousness Bulletin, Fall 1999. P. 8—11.

Churchland P.M. Densmore and Dennett on Virtual Machines and Consciousness // Philosophy and Phenomenological Research. V. 59 (3). 1999. P. 747—767;

— Catching Consciousness in Reccurent Net // Daniel Dennett. Ed. by A. Brook and D. Ross. Cambridge, 2002. P. 64—82.

Clark A. That Special Something: Dennet on Making of Minds and Selves // Dennett's Philosophy: A Comprehensive Assessment. Ed. by D. Ross and A. Brook. MIT Press, Cambridge, 2002. P. 187—205.

Cummings R. What Can be Learned from «Brainstorms»? // Mind, Brain and Function. Essays in Philosophy of Mind. Ed. by J.L Biro and R.W. Shahan. Harvester Press, 1982.

Fodor J. Deconstructing Dennet's Darwin // Fodor J. In Critical Condition. Cambridge MA, 1998. P. 171—178;

— and Leplore E. Dennet: Meaning Holism and the Normativity of International Ascription (and a Little

- More about Davidson) // Fodor J. and Leplore E. Holism. A Shopper's Gaide. Cambridge MA, 1992.
- Foss J. On Evolution of Intentionality as Seen from the Intentional Stance // *Inquiry*. V. 37. № 3. 1994.
- Haugeland J. Pattern and Being // *Dennett and His Critics: Demistifying Mind*. Ed. by Dahlbom B. Oxford, Cambridge, 1993. P. 51—69.
- Hodge J. Daniel Dennett's Dengerous Idea: Evolution and the Meaning of Life // *British Journal for the Philosophy of Science*. V. 48. № 3. 1997.
- Inwagen van P. On D. Dennett's «Elbow Room» // *Nous*. V. 22. № 4. 1988.
- Jackson F. Epiphenomenal Qualia // *Philosophical Quarterly* 32. 1992. P. 127—136.
- Johnson G. What Really Goes On In There // *New York Times*, 10 Nov. 1991.
- 513
- Kline A.D., Robinson W.S. On Brainstorms // *Nature and System*. V. 2. № 1. 1980. P. 7—48.
- Korb B.K. Stage Effects in the Cartesian Theatre: A Review of Daniel Dennett's «Consciousness Explained» // *Psyche*. № 1 (4). Dec. 1993.
- Lukas Book. Towards a Theory of Reflexive Intentional Systems // *Synthese*. V. 118. P. 105—117.
- McGinn C Logic, Mind and Mathematics // *Dennett and His Critics: Demistifying Mind*. Ed. by Dahlbom B. Oxford, Cambridge, 1993;
- Can We Solve the Mind-Body Problem? // *Mind*. V. 98. 1989. P. 349—366.
- Mangan B. Dennett, Consciousness and Sorrows of Functionalism // *Consciousness and Cognition*. V. 2. P. 1—17.
- Nagel T. What We Have in Mind when We Say We are Thinking // *The Wall Street Journal*. Nov. 7. 1991.
- O'Brien G., Opie J. A Defense of Cartesian Materialism // *Philosophy and Phenomenological Research*. V. 59. № 4. 1999. P. 939—963.
- Putnam H. Reflexive Reflections // *Journal of Philosophy*. V. 82. № 1. 1985.
- Reuter Martina. Pre-reflective Intentionality // *Synthese*. V. 118. № 1. 1999. P. 69—88.
- Rorty R. Incorrighibility as the Mark of Mental // *Journal of Philosophy*. V. 67. 1970;
- Dennett on Awareness // *Philosophical Studies*. V. 23. 1972. P. 153—162;
- Functionalism, Machines and Incorrighibility // *Journal of Philosophy*. V. 69. 1972. P. 203—220;
- Comment on Dennett // *Synthese*. V. LUI. 1982. P. 181—187;
- Contemporary Philosophy of Mind // *Synthese*. V. LIII. 1982. P. 323—348;
- 514
- Holism, Intrinsicity and the Ambition of Transcendence // *Dennett and His Critics. Demistifying Mind*. Ed by Bo Dahlbom, Cambridge Mass., 1993. P. 184—202;
- Response to Dennett // *Rorty and his Critics*. Ed. by R.B. Brandom, Oxford, Maiden MA, 2002. P. 101—108.
- Sharpe R.A. Dennett's Journey Towards Panpsychism // *Inquiry*. V. 32. № 2. 1989.
- Stich S.P. Dennet on Intentional Systems // *Philosophical Topics*. V. 12. № 1. 1981.
- Fodor J. Deconstructing Dennett's Darwin // *Mind and Language*. V. 11. 1996.
- Ross D. Dennett's Conceptual Reform // *Behavior and Philosophy*. V. 22. 1994. P. 41—52.
- Thought and Object: Essays on Intentionality. Ed. by Woodfield, Oxford, 1982.
- Voorhees B. Dennett and the Deep Blue Sea // *Journal of Conscious Studies*. № 7. 2000. P. 53—69.
- Wilkes K. V Functionalism, Psychology and Philosophy of Mind // *Philosophical Topics*. V. XII. № 1. 1981.
- Wilson D.C. Functionalism and Moral Personhood: one View Considered // *Philosophy and Phenomenological Research*. V. XLIV. № 4. 1984.
- Yu Paul and Fuller G. A Critique of Dennett // *Synthese*. V. 66. № 3. 1986.

## Общая литература по проблеме сознания на английском языке

Anscombe G.E.M. *Intention*. Oxford, 1963. Armstrong D.M. *A Materialist Theory of Mind*. L., 1968; — and Malcolm N. *Consciousness and Causality*. Oxford, 1984. Baars B. *A Cognitive Theory of Consciousness*. Cambridge, 1988.

515

Bekkermann A., Flohr H. and Kim J. *Emergence and Reduction*. De Gruyter, Berlin, 1992.  
Block N. *What is Functionalism?* // *Readings in Philosophy of Psychology*. V. 1. Cambridge, 1980;  
— *Troubles with Functionalism* // *Perception and Cognition: Issues in the Foundations of Psychology*, *Minnesota Studies in Philosophy of Science*. V. XI. Minneapolis, 1978.  
Bunge M. *The Mind-Body Problem: A Psychological Approach*. Oxford, 1980.  
Bunge M. *Scientific Materialism*. Dordrecht, Boston, L., 1981.  
Block N. and Fodor J. *What a Psychological States are not* // *Philosophical Review*. V. 81. P. 159—181.  
Body, Mind and Method. D. Gustafson and B. Tapscott, eds., Dordrecht, 1979.  
Carnap R. *Psychology in Physical Language* // *Logical Positivism*. Ed. A.J.Ayer, N.Y., 1959.  
Chalmers D. *The Conscious Mind. In Search of a Fundamental Theory*. Oxford, 1996.  
Chomsky N. *New Horizons in the Study of Language and Mind*. Cambridge MA, 2000.  
Churchland P.M. *Scientific Realism and the Plasticity of Mind*. Cambridge, 1979;  
— *Matter and Consciousness*. Cambridge, L., 1984;  
— *A Neurocomputational Perspective: the Nature of Mind and the Structures of Science*. Cambridge MA, 1990;  
— *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. Cambridge, 1986;  
— *The Engine of Reason. The Seat of the Soul*. Cambridge MA, 1995.  
Clark A. *Being There: Putting Brain, Body and World Together Again*. Cambridge MA, 1997;

516

— *Mindware: An Introduction to the Philosophy of Cognitive Science*. Oxford, 2001.  
Conee E. *A Defence of Pain* // *Philosophical Studies*. V. 46. №2. Sept 1984. P. 239—248.  
*Concepts, Theories and Mind-Body Problem* // *Minnesota Studies in the Philosophy of Science*. V. II. Minneapolis, 1952.  
*Consciousness in Modern Science*. Ed. by A. Marcel and E.Bisiach. Oxford, 1988.  
*Contemporary Materialism*. Ed. by P.K. Miser and J.D. Trout. L., 1995.  
Crane T. *Elements of Mind. An Introduction to the Philosophy of Mind*. Oxford, 2001;  
— and Mellor D.H. *There is no Question of Phisycalism* // *Mind*. V. 99. 1990. P. 185—206.  
Crick F. *The Astonishing Hypothesis: the Scientific Search for the Soul*. N.Y., 1994.  
Davidson D. «Mental Events» // *Experience and Theory*. Ed by L. Foster and J. Swanson. Massachusetts, 1970;  
— *Essays on Actions and Events*. Oxford, 1980. Damasio A. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, N.Y., 1999.  
Dawkins R. *The Selfish Gene*. Oxford, 1989 (2<sup>nd</sup> edn).  
*Dimensions of Mind*. Ed. S. Hook. N.Y., 1960.  
Dreyfus H. & Dreyfus S. *Mind over Machine*. N.Y., 1986.  
Ericsson K.A. and Simon H. *Verbal Reports as Data* // *Psychological Review*. V. 87. 1980. P. 215—250.  
Feigl H. *The «Mental» and «Physical»* // *Minnesota Studies in the Philosophy of Science*. V. 3. Ed. by Feigl H. et al., Minneapolis, 1958. P. 370-457.  
Feyerabend P.K. *Materialism and the Mind-Body Problem* // *Realism, Rationalism and Scientific Method. Philosophical Papers*. V. I. Cambridge, L., N.Y., 1981. P. 161—175.

517

Fodor J. *The Language of Thought*. N.Y., 1975;  
— *Representation. Philosophical Essays on the Foundations of Cognitive Science*. Socex, 1984;

- In *Critical Condition*. Boston, 1998.
- Flanagan O. *Consciousness Reconsidered*. Cambridge, 1992;
- *The Science of Mind*. Cambridge MA, 1984.
- Foster J. *The Immaterial Self: a Defence of the Cartesian Dualist Conception of Mind*. L., 1991;
- *The Case for Idealism*. L., 1982.
- Frankfurt H. *Freedom of the Will and the Concept of a Person* // *Journal of Philosophy*. V. LXVIII. № 1. 1971.
- Functionalism and Philosophy of Mind* // *Philosophical Topics*. V. XII. № 1. 1981.
- The Future of Folk Psychology: Intentionality and Cognitive Science*. Ed. by J.D. Greenwood. Cambridge, 1991.
- Howell R. *Fictional Objects: How They are and How they are not* // *Body, Mind and Method*. Ed by D.F. Gustafson and Tapscott B.L. Dordrecht, 1979. P. 241—294.
- Geach P. *Mental Acts*. L. 1957.
- Goldman A.I., Churchland P.M. *Epistemology and Philosophy of Mind* // *Journal of Philosophy*. V. 84. № 10. 1987. P. 537—555.
- Goodman M.F. *Introduction* // *What is a Person*. Ed. by Goodman M.F. Clifton, 1988.
- Gould S.J. *Darwinian Fundamentalism* // *New York Review of Books*, 12.06.97.
- Gunderson K. *Asymmetries and Mind-Body Perplexities* // *Materialism and the Mind-Body Problem*. Englewood Cliffs, 1971;
- *Mentality and Machines*. Doubleday Anchor, 1971. Hofstadter D. *Godel, Escher, Bach, an Eternal Golden Braid*. N.Y., 1979.
- Haugeland J. *Artificial Intelligence: The Very Idea*. Cambridge MA, 1985.
- 518**
- Humphrey N. *A History of Mind*. L., N.Y., 1992. Hurley S. *Consciousness in Action*. Cambridge MA, 1998. Jackson F. *Perception*. Cambridge, 1977;
- *Epiphenomenal Qualia* // *Philosophical Quarterly*. V. 32. 1982. P. 127—136.
- Kenny A. *Metaphysics of Mind*. Oxford, 1989. Kim J. *Supervenience and Mind*. Cambridge, 1993;
- *Philosophy of Mind*. Boulder, 1996;
- *Mind in a Physical World*. Cambridge MA, 1998;
- *The Mind-body Problem: Taking Stock after Forty Years* // *Philosophical Perspectives*, V. 11, *Mind, Causation and World*, 1997.
- Kirk R. *Raw Feeling: A Philosophical Account of the Essence of Consciousness*. Oxford, 1994.
- Knowing our Own Minds*. Ed. By C.Wright, B.C. Smith, C MacDonald. Oxford, 1998.
- Kripke S. *Wittgenstein on Rules and Private Language*. Cambridge MA, 1982.
- Lakoff G. and Johnson M. *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*. N.Y., 1999;
- Leiber J. *An Invitation to Cognitive Science*. Cambridge MA, 1991.
- Levine J. *Materialism and Qualia: The Explanatory Gap* // *Pacific Philosophical Quarterly*. V. 64. P. 354—361.
- Lewis D. *Papers in Metaphysics and Epistemology*. Cambridge, 1999.
- Lockwood M. *Mind, Brain and the Quantum: the Compound I*. Oxford, 1991.
- Lycan W.G. *Consciousness and Experience*. Cambridge MA, 1996.
- Materialism and the Mind-Body Problem*. Ed. by Rosenthal, Englewood Cliffs (N.J.), 1971.
- McGinn C *The Character of Mind*. Oxford, N.Y., 1982;
- 519**
- *Problem of Consciousness*. Oxford, 1990.
- The Mind-Brain Identity Theory*. Ed. by C.V. Borst, L, N.Y, 1970.
- Mind, Language and Reality: Philosophical Papers*. V. 2, Ed. by H. Putnam. Cambridge, 1980.
- Mind, Brain and Function. Essays in Philosophy of Mind*. Ed. by J.L Biro and R.W. Shahan. Harvester Press, 1982.
- Mind, Matter and Method* // Ed. by P. Feyerabend, C Maxwell. Minneapolis, 1966.

- Nagel T. Physicalism // *Philosophical Review*. V. LXXIV, 1965. P. 339—356;  
— What is it like to be a Bat // *Philosophical Review*. V. LXXXIII, 1974. P. 435—445;  
— *Mortal Questions*. Cambridge, 1979;  
— *The View from Nowhere*. Oxford, 1986;  
— What we have in mind when we say we are thinking // *The Wall Street Journal*. № 7. 1991;  
— *The Last Word*. N.Y., Oxford, 1997.
- Nisbett R. and Wilson T.D.C. Telling More than We Know: Verbal Reports on Mental Processes // *Psychological Review*. V. 84. 1977. P. 231—259.
- Objections to Physicalism. Ed. by H. Robinson. Oxford, 1993.
- Papineau D. *Philosophical Naturalism*. Oxford, 1993.
- Parfit D. Personal Identity // *Philosophical Review*. V. 80. 1971. P. 3—27.
- Penrose R. *The Emperor's New Mind: Concerning Computers, Minds and the Laws of Physics*. Oxford, 1985.
- Pettit Ph. A Definition of Physicalism // *Analysis*. V. 53. 1993. P. 3—21.
- Petitot J., Varela F., Roy J.M. *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*. Stanford CA, 1999.
- 520
- Philosophy and Connectionist Theory*. Ed. by Ramsey W., Stich S., Garon J.L., Erlbaum Assoc, 1991.
- Poland J. *Physicalism: The Philosophical Foundations*. Oxford, 1994.
- Popper K., Eccles J. *The Self and its Brain*. In *Defence of Interactionism*. B., N.Y., L., 1977.
- Popper K. *Knowledge and Body-Mind Problem*. In *Defence of Interactionism*. L., N.Y., 1994.
- Reed E. Descartes's Corporeal Ideas Hypotheses and the Origin of Scientific Psychology // *The Review of Metaphysics*. V. XXXV. № 4. 1982.
- Robinson H. *Matter and Sense. Critique of Contemporary Materialism*. Cambridge and N.Y., 1982.
- Rorty A.O. *Mind in Action: Essays in Philosophy of Mind*. Boston, 1988.
- Readings in the Philosophy of Psychology*. Ed. by N. Block, 2 vols. L., 1980.
- Quine V.O. *Epistemology Naturalised // Ontological relativity and Other Essays*. Cambridge MA, 1969.
- Sellars W. *Empiricism and the Philosophy of Mind // The Foundations of Science and the Concept of Psychology and Psychoanalysis*. Minneapolis, 1956.
- Siewert Ch. P. *The Significance of Consciousness*. Princeton, 1998.
- Skinner B.F. *Science and Human Behavior*. N.Y., 1963;  
— *Beyond Freedom and Dignity*. N.Y., 1971.
- Searle J.R. *Minds, Brains and Science*. Cambridge MA, 1984;  
— *The Rediscovery of Consciousness*. Cambridge MA, 1992;  
— *The Construction of Social Reality*. N.Y., 1995;  
— *Mystery of Consciousness*. N.Y., 1998.
- Smart J.J.C. *Philosophy and Scientific Realism*. L., Atlantic Highlands, 1963.
- 521
- Stich S. *From Folk Psychology to Cognitive Science*. Cambridge MA., 1983.
- Thalberg I. Immateriality // *Mind*. V. XIII. 1983.
- Thought and Object: Essays on Intentionality*. Ed by A. Woodfield. Oxford, 1981.
- Towards a Science of Consciousness II: The Second Tucson Discussions and Debates*. Cambridge MA, 1998.
- Turing A. *Computing Machinery and Intelligence // Mind*. V. 59. P. 433-460.
- Wasserman G.D. *Materialism and Mentality // The Review of Metaphysics*. V. 35. № 4. 1982.
- Wegner D. *The Illusion of Conscious Will*. Cambridge MA, 2002.
- Wilkes K.V. *Physicalism*. L., 1978.
- Wilson Ed.O. *Sociobiology: the New Syntheses*. Cambridge MA, 1975.
- Wilson E. *The Mental as Physical*. L., 1979.
- Winograd T. *Understanding Natural Language*. N.Y., 1972.



Wright R. Nonzero: the Logic of Human Destiny. N.Y., 2000.

## Общая литература о сознании на русском языке

- Абрамова Н.Т. Несловесное мышление. М., 2002.
- Аналитическая философия. Избранные тексты. М., 1993.
- Аналитическая философия: становление и развитие. Под ред. А.Ф. Грязнова. Пер. с англ. (колл). М., 1998.
- Арбиб М. Метафорический мозг. М., 1976.
- Балмаева С.Д. Методология исследования сознания в буржуазной англоязычной философии XX века: «аналитическая традиция» // Научный метод и методологическое сознание. Свердловск, 1986. С. 140—149.
- Баксанский О.Е., Кучер Е.Н. Когнитивные науки: междисциплинарный подход. М., 2003.
- 522
- Блинов А.Л. Формальная онтология действий // Логический анализ языка: модели действия. М., 1992.
- Богданов А.А. Эмпириомонизм. М., 2003.
- Боррадори Дж. Американский философ. Пер. с англ. (колл.), М., 1999.
- Братко А.А., Кочергин А.Н. Информация и психика. М., 1977.
- Брентано Ф. Избранные труды. Пер. с нем. В.В. Анашвили. М., 1996. Т. 1.
- Брушлинский А.В. Проблемы психологии субъекта. М., 1994.
- Бунге М. Несостоятельность психофизического дуализма // Философские науки. № 2. 1979.
- Величковский Б.М. Современная когнитивная психология. М., 1982.
- Винер Н. Творец и робот. Пер. с англ. Г.Н. Поворова. М., 1966.
- Витгенштейн Л. Логико-философский трактат. Пер. с англ. Г.Н. Добронравова. М., 1958.
- Витгенштейн Л. Философские исследования. Пер. с нем. М.С.Козловой, Ю.А.Асеева. М., 1996.
- Выгодский Л.С. Исторический смысл психологического кризиса // Выгодский Л.С. Собр. соч. М., 1982. Т. 1.
- Гальперин П.Я. Введение в психологию. М., 1976.
- Грязнов А.Ф. Язык и деятельность: критический анализ витгенштейнианства. М., 1991.
- Гудмен Н. Способы создания миров. Пер. с англ. А.Л. Никифорова, Е.Е. Ледникова, М.В. Лебедевой, Т.А. Дмитриева. М., 2001.
- Гуссерль Э. Собрание сочинений. Т. I—IV. Пер. с нем. В.И.Молчанова. М., 1994—2001.
- Дегутис А.Ю. Язык, мышление и действительность. Вильнюс, 1984.
- 523
- Дегутис А.Ю. Онтологическая проблематика в современной аналитической философии // Проблемы онтологии в современной буржуазной философии. Под ред. Т.А. Кузьминой. Рига, 1988. С. 205—225.
- Джемс У. Существует ли «сознание»? // Новые идеи в философии. Сб. 4. СПб., 1913.
- Докинс Р. Эгоистичный ген. М., 1993.
- Дрейфус Х. Чего не могут вычислительные машины. М., 1978.
- Дубровский Д.И. Информация. Сознание. Мозг. М., 1980;
- Проблема идеального. Субъективная реальность. 2-е изд. М, 2002;
- Проблема духа и тела: возможности решения (В связи со статьей Т. Нагеля «Мыслимость невозможного и проблема духа и тела») // Вопросы философии. 2002. № 10;
- Новое открытие сознания? (По поводу книги Джона Сёрля «Открывая сознание заново» // Вопросы философии. 2003. № 7.
- Ильенков Э.В. Идеальное // Философская энциклопедия. М., 1962. С. 219—227.
- Карнап Р. Значение и необходимость. М., 1958.
- Козлова М.С. Философия и язык. М., 1972.
- Концепция виртуальных миров и научное познание. Ред. И.А. Акчурин, С.Н. Коняев. СПб., 2000.

- Куайн У.О. Слово и объект. Пер. с англ. А.З. Черняка, Т.А. Дмитриева. М., 2000.
- Ламетри Ж. Человек-машина // Ламетри Ж. Сочинения. М., 1989.
- Лекторский В.А. Деятельностный подход: смерть или возрождение // Вопросы философии. 2001. № 2;
- Сознание // Новая философская энциклопедия. М., 2001. С. 589—591;
- 524
- Садовский В.Н. Основные идеи «генетической эпистемологии» Жана Пиаже // Вопросы философии. 1964. № 4.
- Леонтьев А.Н. Очерк развития психики. М., 1947.
- Макеева Л.Б. Философия Х.Патнэма. М., 1996.
- Максимов Л.В. Когнитивизм как парадигма гуманитарно-философской мысли. М., 2003.
- Марголис Д. Личность и сознание. Пер. с англ. А.Ф. Грязнова, Д.И. Дубровского. М., 1986; Там же: Дубровский Д.И. О книге Дж. Марголиса и его концепции «эмерджентистского материализма». С. 5—30.
- Меркулов И.П. Когнитивная эволюция. М., 1999;
- Информационная природа сознания // Полигнозис. 2000. № 4;
- Когнитивная модель сознания // Эволюция. Мышление. Сознание. Под ред. И.П. Меркулова. М., 2004.
- Минский М. На пути к созданию искусственного разума // Вычислительные машины и мышление. М., 1967.
- Мотрошилова Н.В. «Идеи I» Эдмунда Гуссерля как введение в феноменологию. М., 2003.
- Михайлов Ф.Т. Загадка человеческого Я. М., 1976.
- Мышление и язык. Под ред. Д.П. Горского. М., 1957.
- Нагель Т. Мыслимость невозможного и проблема духа и тела // Вопросы философии. 2002. № 8.
- Найссер У. Познание и реальность. Под ред. Б.М. Величковского. Пер. с англ. В.В. Лучкова. М., 1981.
- Никитин Е.П. Объяснение — функция науки. М., 1970.
- Огурцов А. Проблема универсалий в философии XX века // Истина и благо: универсальное и сингулярное. М., 2002. С. 260—339.
- Остин Д. Избранное. Пер. с англ. Л.Б. Макеевой, В.П. Руднева. М., 2002.
- 525
- Пассмор Дж. Современные философы. Пер. с англ. Л.Б. Макеевой. М., 2002.
- Патнэм Х. Философия сознания. Пер. с англ. Л.Б. Макеевой, О.А. Назаровой, А.Л. Никифорова. М., 1998.
- Пиаже Ж. Избранные психологические труды. М., 1969.
- Полани М. Личностное знание. На пути к посткритической философии. Под ред. В.А. Лекторского, В.А. Аршинова. Пер. с англ. М.Б. Гнедовского, Н.М. Смирновой, Б.А. Старостина. М., 1995.
- Поппер К. Объективное знание. Под ред. В.Н. Садовского. Пер. с англ. Д.Г. Лахути. М., 2002.
- Порус В.Н. Рациональность. Наука. Культура. М., 2002.
- Прист С. Теории сознания. Пер. с англ. А.Ф. Грязнова. М., 2000.
- Проблема сознания в современной западной философии. Под ред. Т.А. Кузьминой. М., 1989.
- Пушкин В.Г. К проблеме «искусственного интеллекта» // Вопросы философии. 1978. № 5;
- Проблема сознания и кибернетика // Философские науки. 1981. № 3.
- Райл Дж. Понятие сознания. Под ред. В.П. Филатова. Пер. с англ. М., 2000.
- Рассел Б. Исследования значения и истины. Пер. с англ. Е.Е. Ледникова, А.Л. Никифорова. М., 1999.
- Родин С.Н. Идея коэволюции. Новосибирск, 1991.
- Рациональность на перепутье. Кн. 1. Под ред. В.А. Лекторского. М., 1999.
- Рорти Р. Философия и зеркало природы. Пер. с англ. В.В. Целищева. Новосибирск, 1997.
- Рорти Р. Случайность, ирония и солидарность. Пер. с англ. И. Хестановой, Р. Хестанова. М., 1996.

Рубинштейн СЛ. Бытие и сознание. М., 1957.

526

Садовский В.Н. История создания, теоретические основы и судьба эмпириомонизма А.А. Богданова // Богданов А.А. Эмпириомонизм. М., 2003. С. 340—365.

Самсонов В.Ф. О философии Уилларда Куайна // Самсонов В.Ф. О философии. Современные дискуссионные проблемы. Челябинск, 2002.

Саймон Г. Наука об искусственном. Пер. с англ. Э.Л. Напельбаума. М., 1972.

Селларс У. Научный реализм или «миролюбивый» инструментализм // Структура и развитие науки. Под ред. Б. Грязнова и В.Н. Садовского. М., 1978. С. 353—395.

Селиванов Ю.Р. Сознание масс-медиа. М., 2002.

Сёрль Дж. Открывая сознание заново. Пер. с англ. А.Ф. Грязнова. М., 2002.

Смарт Дж.Дж. Конфликтующие точки зрения по проблеме объяснения // Структура и развитие науки. Под ред. Б. Грязнова и В.Н. Садовского. М., 1978. С. 337—352.

Сперри Р.У. Перспективы менталистской революции и возникновение нового научного мировоззрения // Мозг и разум. М., 1994.

Спиркин А.Г. Сознание и самосознание. М., 1972.

Степин В.С. Теоретическое знание. М., 2000.

Тьюринг А. Может ли материя мыслить? Пер. с англ. М., 1960.

Урсул А.Д. Философия и общенаучный характер проблемы искусственного интеллекта // Вопросы философии. 1979. № 4.

Фрейд З. Психология бессознательного. М., 1989.

Хилл Т.И. Современные теории познания. Пер. с англ. М., 1965.

Швырев В.А. О понятиях «открытой» и «закрытой» рациональности // Рациональность на перепутье. Кн. 1. М., 1999.

Шрейдер Ю.А. Присущ ли машине разум? // Вопросы философии. 1975. № 2.

527

Эволюционная эпистемология и логика социальных наук. Карл Поппер и его критики. Под ред. В.Н. Садовского. Пер. с англ. Д.Г. Лахути. М., 2000.

Юлина Н.С. Проблема метафизики в американской философии XX века. М., 1978;

— К. Поппер: мир предрасположенностей и активность самости (аргумент против физикализма) // Философские исследования. М., 1997. № 4;

— Очерки по философии в США. XX век. М., 1999;

— О книге Джона Сёрля «Открывая сознание заново» // Вопросы философии. 2003. № 3. С. 186—191;

— «Открытая Вселенная» и активная самость // На пути к открытому обществу. Идеи Карла Поппера и современная Россия. М., 1998. С. 91—112.

## Указатель имен\*

Абельсон Р. (Abelson R.) 311  
 Айер А. Дж. (Ayer A.J.) 248  
 Апдайк Дж. (Updike J.) 437, 438, 446  
 Арбиб М. (Arbib M.) 21  
 Аристотель (Aristotle) 53, 223, 332, 336, 422, 432, 451  
 Армстронг Д.М. (Armstrong D.M.) 37, 111  
 Баарс Б. (Baars B.J.) 129, 290, 394  
 Барон-Козен С. (Baron-Kohen C.) 290  
 Бергсон А. (Bergson H.) 332, 336  
 Беркли Дж. (Berkley G.) 62, 66  
 Блок Н. (Block N.) 18, 21, 46, 63  
 Богданов А.А. 386  
 Борхес Х.Л. (Borges J.L.) 410, 411, 427  
 Брентано Ф. (Brentano F.) 24, 49, 50, 104, 121, 202, 203, 204, 205, 208, 214, 239, 241  
 Валентинов Н.В. 386 Валери П. (Valéry P.) 398, 427  
 Вейценбаум Дж. (Weizenbaum J.) 311  
 Виноград Т. (Vinograd T.) 311

\* Без имен, упомянутых в библиографии.

529

Витгенштейн Л. (Wittgenstein L.) 10, 18, 21, 33, 34, 36, 59, 60, 61, 62, 82, 86, 165, 188, 203, 223, 285, 287, 288, 326, 329, 330, 331, 332, 334, 338, 342, 345, 353, 355, 356, 384, 390, 391, 394  
 Гадамер Г.-Г. (Gadamer H.-G.) 327, 340, 347  
 Гастон Ф. (Guston Ph.) 426, 427  
 Гегель Г.В.Ф. (Hegel G.W.F.) 12  
 Гемпель К. (Hempel K.) 32, 33  
 Гоулд С. (Gould S.) 21, 166, 233  
 Грайс Х. (Grice H.) 230, 231, 248, 452, 464, 465, 466, 468, 469  
 Грегори Р. (Gregory R.) 227  
 Гриффин Р. (Griffin R.) 290  
 Гудмен Н. (Goodman N.) 59, 328, 330, 332, 340, 344, 345, 347, 355  
 Гудмен Э. (Goodman A.) 132  
 Гундерсон К. (Gunderson K.) 88, 89  
 Гуссерль Э. (Husserl E.) 50, 205  
 Гэзанига М. (Gazzaniga M.) 440, 441  
 Дальбом Б. (Dahlbom B.) 219, 220, 224, 225, 288, 289  
 Дарвин Ч. (Darwin Ch.) 7, 11, 48, 147, 148, 163, 164, 165, 170, 187, 188, 190, 191, 192, 398, 400, 401, 405, 421, 422, 423, 427  
 Декарт Р. (Descartes R.) 5, 12, 23, 24, 31, 49, 50, 52, 54, 56, 87, 123, 174, 175, 202, 286, 353, 427  
 Деннет С. (Dennett S.) 427  
 Деррида Ж. (Derrida J.) 288, 324, 327, 328, 340  
 Джейнс Дж. (Janes J.) 441

530

Джексон Фр. (Jackson Fr.) 78, 80  
 Джеймс У. (James W.) 26, 37, 155, 174, 178, 297, 332, 335, 387  
 Докинс Р. (Dawkins R.) 11, 165, 166, 167, 172, 224, 327, 421, 427  
 Достоевский Ф.М. 437  
 Дучэмп М. (Duchamp M.) 424, 427  
 Дьюи Дж. (Dewey J.) 9, 335, 340, 387  
 Дэвидсон Д. (Davidson D.) 38, 204, 236, 264, 289, 295, 325, 332, 333, 334, 337, 338  
 Кармилоф-Смит Э. (Karmiloff-Smith A.) 94  
 Кант И. (Kant I.) 12, 246, 261, 329, 382, 451  
 Карнап Р. (Carnap R.) 32, 33, 38, 354, 366, 391, 392  
 Каспаров Г. 193, 194, 195, 417, 418, 419, 420  
 Кестлер А. (Koestler A.) 56

Ким Дж. (Kim J.) 41, 295  
 Кингсборн М. (Kinsbourne M.) 290, 394  
 Кларк Ст. (Clark St.) 21  
 Кларк А. (Clark A.) 171, 172, 173, 295, 373  
 Конан-Дойль А. (Conan-Doyle A.) 431  
 Коуп Д. (Cope D.) 193, 416, 419, 420, 427  
 Коэн Л.Дж. (Cohen L.J.) 87  
 Крипке С. (Kripke S.) 46, 204, 236  
 Куайн У.О. (Quine W.O.) 6, 21, 35, 36, 59, 203, 204, 208, 209, 217, 236, 245, 254, 260, 276, 281, 283, 287, 288, 295, 328, 342, 366, 374, 388, 392, 397  
 Кун Т. (Kuhn Th.) 345, 347, 379  
 Кэмпбелл М. (Campbell M.) 193, 417, 420  
 531  
 Лакатос И. (Lacatos I.) 242  
 Ламетри Ж. (La Mettrie J.) 261, 277  
 Левайн Дж. (Levine J.) 123  
 Левельт П. (Levelt P.) 139  
 Левонтин Р. (Lewontin R.C.) 166  
 Лейбер Дж. (Leiber J.) 165  
 Лейбниц Г. (Leibniz G.) 332, 336  
 Ленин В.И. 43  
 Леплор Е. (Leplore E.) 237, 238, 239  
 Либет А. (Libet A.) 132  
 Локвуд М. (Lockwood M.) 46, 49, 56, 65, 100, 134, 295  
 Локк Дж. (Locke J.) 24, 66, 71, 72, 166, 449  
 Льюис Д. (Lewis D.) 33  
 МакГинн К. (McGinn C.) 21, 26, 46, 93, 97, 123, 142,  
 Макгир В. (McGeer V.) 427  
 143, 288, 295, 374  
 МакКей Э. (MacKey A.) 269, 270  
 МакКензи Р. (MacKenzie R.) 400, 407, 427  
 Макинтайр А. (MacIntyre A.) 265, 269  
 Малколм Н. (Malcolm N.) 265, 266  
 Маркс К. (Marx K.) 246, 257  
 Марсель Т. (Marcel T.) 290, 394, 427  
 Мах Э. (Mach E.) 340  
 Мелвилл Г. (Melville G.) 431, 433  
 Милликэн Р. (Millikan R.) 204, 236  
 Минский М. (Minsky M.) 164, 204, 236, 318  
 532  
 Нагель Т. (Nagel Th.) 21, 46, 65, 87, 93, 123, 142, 143, 204, 248, 289, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 306, 310, 324, 325, 331, 332, 333, 334, 362, 363  
 Нейман Дж. (Neumann J.) 149, 150, 154, 155, 157, 158, 169, 173  
 Нейрат О. (Neurath O.) 292  
 Ницше Ф. (Nietzsche F.) 257, 288, 326, 328  
 Ньютон И. (Newton I.) 178, 330, 375, 428  
 Ньюэлл Э. (Newell A.) 10, 116  
 Оруэлл Дж. (Orwell G.) 134 Остин Дж.Л. (Austin J.L.) 394  
 Парфит Д. (Parfit D.) 182  
 Патнэм Х. (Putnam H.) 10, 21, 41, 112, 116, 117, 118, 119, 120, 122, 269, 295, 325, 347, 452  
 Паундстон У. (Poundstone W.) 404, 408, 427  
 Пенроуз Р. (Penrose R.) 21, 46, 132, 198  
 Перри Р.Б. (Perry R.B.) 387  
 Пиаже Ж. (Piaget J.) 242  
 Пикассо П. (Picasso P.) 407, 424, 427  
 Платон 23, 199, 355, 358, 423, 481, 489  
 Плейс Н. (Place N.) 37, 111  
 Поппер К.Р. (Popper K.R.) 21, 46, 56, 163, 199, 242, 243, 278, 296, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 385



Проктор Р. (Proctor R.) 490  
 533  
 Райл Дж. (Ryle G.) 6, 10, 21, 33, 34, 36, 56, 113, 117, 188, 203, 287, 288, 289, 328, 330, 331, 332, 334, 355, 387, 391, 394, 443  
 Рассел Б. (Russell B.) 285, 387, 390  
 Рейхенбах Г. (Reichenbach H.) 179, 391, 392, 428, 432  
 Розенталь Д. 295  
 Ройс Дж. (Royce J.) 339  
 Рорти А.О. (Rorty A.O.) 248  
 Рорти Р. (Rorty R.) 5, 9, 13, 21, 38, 39, 67, 103, 204, 219, 224, 236, 295, 296, 297, 304, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 358, 359, 360, 361, 362, 363, 384, 391, 393, 452, 483, 484, 485  
 Росс Д. (Ross D.) 285, 290  
 Роулз Дж. (Rowls J.) 451, 470, 471, 472  
 Роэ Л.М. (Rohe L.M.) 415, 427  
 Саймон Г. (Simon H.) 10, 116, 152, 220, 225  
 Сартр Ж.-П. (Sartre J.-P.) 453  
 Селларс У. (Sellars W.) 21, 59, 185, 204, 236, 264, 265, 277, 295, 328, 334, 343, 394; 542  
 Сёрль Дж.Р. (Searle J.R.) 18, 21, 26, 46, 64, 65, 93, 123, 154, 191, 192, 204, 219, 222, 236, 294, 295, 305, 306, 307, 308, 311, 312, 313, 314, 315, 316, 319, 320, 321, 322, 323, 324, 325, 331, 332, 333 334, 343, 362, 363, 400, 401, 402, 427  
 Скиннер Б. (Skinner B.) 21, 32, 202, 208, 225, 226, 248, 249, 250, 254, 295, 448  
 534  
 Сمارт Дж.Дж. (Smart J.J.) 36, 37, 111  
 Сократ 489  
 Стич С. (Stich S.) 295  
 Стросон П. (Strawson P.) 248, 265, 269, 270, 277, 451, 452  
 Тейлор Ч. (Taylor Ch.) 205  
 Тулмин С. (Toulmin S.) 242  
 Тьюринг А.М. (Turing A.M.) 10, 11, 12, 40, 41, 43, 48, 110, 111, 116, 117, 118, 119, 120, 121, 147, 148, 149, 150, 156, 162, 164, 165, 170, 200, 205, 286, 313, 316, 319, 372, 381, 389  
 Уайтхед А.Н. (Whitehead A.N.) 42, 332  
 Уилкс Й. (Wilks Y.) 305  
 Уилсон Э.О. (Wilson Ed.O.) 11, 165, 327, 363, 479  
 Уилсон Д. (Wilson J.D.) 259, 260  
 Уильямс Б. (Williams B.) 484  
 Фейгл Г. (Feigl H.) 36, 37, 111, 366  
 Фейерабенд П. (Feyerabend P.) 38, 325, 379  
 Фишер Б. (Fischer B.) 194, 418, 427  
 Фланеган О. (Flanagan O.) 46  
 Флю Э. (Flew A.) 265, 452  
 Фодор Дж. (Fodor J.) 18, 21, 93, 97, 123, 174, 204, 222, 235, 236, 237, 238, 239, 240, 241, 242, 243, 289, 294, 295, 325, 355  
 Фостер Дж. (Foster J.) 46, 56, 93, 286, 295  
 Франкфурт Х. (Frankfurt H.) 248, 453, 454, 473, 474, 475  
 535  
 Хайдеггер М. (Heidegger M.) 288, 324, 326, 328, 340  
 Харман Дж. (Harman J.) 295  
 Харнад СР. (Harnad S.R.) 295  
 Хемпшайр С. (Hampshire C.) 277  
 Хинтикка Я. (Hintikka J.) 5, 451, 471  
 Хогелэнд Дж. (Haugeland J.) 204, 219, 224, 236, 288  
 Хомский Н. (Chomsky N.) 21, 35, 93, 97, 138, 355  
 Хосперс Дж. (Hospers J.) 266  
 Хофстадтер Д. (Hofstadter D.) 204, 236, 295, 415, 427

Хэмфри Н. (Humphrey N.) 76, 427  
Черчленд П. (Churchland P.) 18, 21, 38, 100, 147, 157, 172, 173, 204, 236, 242, 263, 289, 295, 296, 325, 373  
Черчленд П. (Churchland P.) 295  
Чини Д. (Chiney D.) 290  
Чэлмерс Д.Дж. (Chalmers D.J.) 18, 21, 38, 100, 147, 157, 172, 173, 204, 236, 242, 263, 289, 295, 296, 325, 373  
Шекспир У. (Shakespeare W.) 190, 404, 427  
Шелли М. (Shelley M.) 427  
Шлик М. (Shlick M.) 363, 391, 392  
Шопенгауэр А. (Schopenhauer A.) 24  
Шумейкер С. (Shoemaker S.) 185  
Эдисон Т. (Edison Th.) 427  
Эйнштейн А. (Einstein A.) 372, 375  
Экинс К. (Akins K.) 76  
536  
Экклз Дж. (Eccles J.) 46, 56, 132, 364, 365, 367, 376  
Энскомб ГЭ. (Enscombe G.E.) 453, 454, 463, 467, 468, 472, 475  
Юлина Н.С. 397  
Юм Д. (Hume D.) 11, 24, 26, 66, 166, 175, 175, 466  
Юшкевич П.С. 386  
537

## Содержание

### ЧАСТЬ I.

#### ГОЛОВОЛОМКИ ПРОБЛЕМЫ СОЗНАНИЯ.....5

<b>I. Полвека атак на «бастион сознания».....</b>	<b>5</b>
1. О Дэниеле Деннете.....	5
2. Презентация философии Деннета: трудности ее исследования.....	14
3. Тайна феномена сознания.....	23
4. История атак на дуализм.....	31
<b>II. Деконструкция опыта и сознания.....</b>	<b>49</b>
1. Иллюзии «Картезианского Театра» .....	49
2. «Картезианский материализм» .....	55
3. Миражи «феноменологического сада».....	59
4. «Дисквалификация квалиа».....	69
5. Иметь опыт — значит иметь суждение об опыте.....	80
6. Почему нельзя создать чувствующий боль компьютер? .....	84
7. Животные и младенцы, не имеющие языка, не имеют сознания.....	92
8. Гетерофеноменология или как исследовать сознание эмпирически.....	101
<b>III. Сознание в рамках «Модели Множественных набросков».....</b>	<b>110</b>
1. От теории тождества к функционализму «Машины Тьюринга».....	110
2. «Модель Множественных набросков»: децентрализация сознания....	125
3. Аномалии восприятия времени: «Сталинская» и «Оруэлловская» модели его редактирования.....	132
4. Поиск значений: обратная связь невербального и вербального.....	137
<b>IV. «Машина Тьюринга», «Машина Дарвина» и виртуальная «Машина Сознания».....</b>	<b>147</b>
1. Компьютерные метафоры и философия.....	147
2. Виртуальная «Машина Сознания»: ее «размещение» в органике мозга.....	153
3. Эволюционизм, язык и роль меме-вирусов в производстве сознания.....	162
<b>V. Самость как «центр нарративной гравитации» или почему возможны самостные компьютеры.....</b>	<b>174</b>
1. Биологическая самость и нарративная самость.....	174
2. Почему возможно помыслить самостный компьютер.....	182
3. Свобода воли и детерминизм.....	186
4. Эволюционные алгоритмы и творчество.....	190
5. Где находится мое Я? Телепортация и бессмертие.....	195
<b>VI. Интенциональность как установка рациональности.....</b>	<b>200</b>
1. Бихевиоризм и интенциональность.....	200
2. Физическая установка, установка дизайнера и интенциональная установка.....	206
3. Интенциональность, эволюционный аргумент и «свободно плавающая рациональность».....	214
4. Уровни интенциональности	

538

	и интеллектуальности.....	227
539	5. Первичная и производная интенциональность. Мать-Природа как первичный дизайнер.....	231
	6. Джерри Фодор о телеологизме эволюционного аргумента Деннета.....	235
	<b>VII. Об условиях личностного.....</b>	<b>246</b>
	1. Необходимые и достаточные условия понятия «личностное».....	246
	2. «Личностное» в свете интенциональной установки ..	253
	3. Нормативно личностное и метафизически личностное.....	257
	<b>VIII. О понятии «ответственность» в свете механицистского объяснения человека.....</b>	<b>261</b>
	1. Можно ли включить целевое объяснение человека в физикалистскую картину мира?.....	261
	2. Совместимость механицизма и ответственности в свете теории установок.....	265
	3. Действия, манипуляция и ответственность: каузальные и рациональные факторы.....	272
	<b>ЧАСТЬ II. ДЕННЕТ И КРИТИКА.....</b>	<b>285</b>
	<b>I. Споры вокруг философии Деннета.....</b>	<b>285</b>
	<b>II. Нагель versus Деннет: «объективная феноменология» против панобъективизма.....</b>	<b>297</b>
	<b>III. Сёрль versus Деннет: «Аргумент Китайской Комнаты».....</b>	<b>305</b>
	1. Материализм и «онтология субъективного».....	305
	2. Когнитивизм компьютера и человеческое понимание. Синтаксис и семантика.....	311
540	<b>IV. Деннет versus Рорти: метафоризм, реализм и «базисный инстинкт».....</b>	<b>324</b>
	<i>Рорти о Деннете</i> .....	329
	1. «Коперниканский переворот» в философии сознания: от эссенциализма к прагматизму.....	329
	2. Деннет «сидит на заборе» между реализмом и ирреализмом.....	334
	<i>Деннет о Рорти</i> .....	340
	3. О «привилегированных» репрезентациях старого когнитивизма и «непривилегированных» репрезентациях нового когнитивизма.....	340
	4. Метафоризм, базисный инстинкт и «вегетарианская» концепция истины.....	347
	5. Регулятивный идеал философии: объективность или солидарность?.....	352
	6. О вирусе постмодернизма и «западной» науке.....	356
	<b>V. К. Поппер и Д. Деннет: архитектура сознания согласно «Открытой» и «Закрытой» Вселенной.....</b>	<b>364</b>
	<b>Заключение. Заметки историка философии.....</b>	<b>377</b>
	<b>Приложение 1. Тексты Д.К. Деннета (Перевод с английского Н.С. Юлиной).....</b>	<b>397</b>
	1. По следу Дарвина. Где есмь Я?.....	397
	2. Почему каждый из нас является новеллистом.....	42 8
	3. Условия личностного.....	447
	4. Почему нам важно понимать это правильно. Постмодернизм и истина.....	477

<b>Приложение 2. Библиография.....</b>	<b>493</b>
<b>Указатель имен.....</b>	<b>529</b>
541	



## Contents

<b>PART I. THE «BRAINSTORMS» IN PHILOSOPHY OF MIND.....</b>	<b>5</b>
<b>I. Half-century Attacks on «Bastion of Consciousness».....</b>	<b>5</b>
<b>II. Deconstruction of Experience and Consciousness.....</b>	<b>49</b>
<b>III. Consciousness in the Frame of «The Multiple Draft Model».....</b>	<b>110</b>
<b>IV. «Darwin's Machine», «Turing's Machine» and Virtual «Machine of Consciousness».....</b>	<b>147</b>
<b>V. The Self as a «Center of Narrative Gravity» or Why Computer Could Have a Self.....</b>	<b>174</b>
<b>VI. Intentionality as a Rational Stance.....</b>	<b>200</b>
<b>VII. Conditions of Personhood.....</b>	<b>246</b>
<b>VIII. On «Responsibility» in View of Mechanistic Explanation of Human Being.....</b>	<b>261</b>
<b>PART II. DENNETT AND HIS CRITICS</b>	
<b>I. Discussions around Dennett's Philosophy.....</b>	<b>285</b>
542	
<b>II. Nagel vs. Dennett: «Objective Phenomenology» or Panobjectivism.....</b>	<b>297</b>
<b>III. Searle vs. Dennett: «The Chinese Room Argument».....</b>	<b>305</b>
<b>IV. Dennett vs. Rorty: Metaphorism, Realism, and «Basic Instinct» .....</b>	<b>324</b>
<b>V. K. Popper and D. Dennett: Architecture of Consciousness According to «Open» and «Closed» Universe.....</b>	<b>364</b>
Conclusion. Some Notes by the Historian of Philosophy.....	377
Supplement 1. D.C. Dennett: Selected Papers (Translated by N.S.Yulina).....	397
1. In Darwin's Wake. Where am I?.....	397
2. Why Everyone is a Novelist.....	428
3. Conditions of Personhood.....	447
4. Why Getting It Right Matters. Postmodernism and Truth.....	477
<b>Supplement 2. Bibliography.....</b>	<b>493</b>
<b>Supplement 3. Index of Names.....</b>	<b>529</b>
543	

в розницу можно приобрести в магазине «Пресса для всех» по адресу: г. Москва, Б. Сухаревская площадь, д. 16/18.  
Тел.207-51-13. E-mail: [bozhkoyra@mtu-het.ru](mailto:bozhkoyra@mtu-het.ru)

*Научное издание*

Головоломки проблемы сознания

Концепция Дэниела Деннета

Ответственный за выпуск *Божко Ю. В.*

Подписано в печать с готовых диапозитивов 11.10.2004.

Формат 84 x 108 1/32. Гарнитура тип Таймс. Печать офсетная.

Усл. печ. л. 28,56. Тираж 1000 экз. Заказ 2673.

Издательство «Канон +» *ОН* «Реабилитация». 103051, Москва, Б. Сухаревская пл., д. 16/18.

Республиканское унитарное предприятие

«Издательство «Белорусский Дом печати».

220013, Минск, пр. Ф. Скорины, 79.

ISBN 5-88373-224-0

---

Сканирование и форматирование: Янко Слава (Библиотека Fort/Da) || [slavaaa@yandex.ru](mailto:slavaaa@yandex.ru) ||

[yanko\\_slava@yahoo.com](mailto:yanko_slava@yahoo.com) || <http://yanko.lib.ru> || Исq# 75088656 || Библиотека:

<http://yanko.lib.ru/gum.html> ||

update 09.05.06

---

## Приложения www = SL

### Тексты Деннета из www

#### In Darwin's Wake, Where am I? - December 29, 2000

<http://ase.tufts.edu/cogstud/papers/apapresadd.htm>

APA Presidential Address.

December 29, 2000

Daniel C. Dennett

#### In Darwin's Wake, Where am I?

Philosophy has lost a giant this week. Willard Van Orman Quine died on Christmas Day at the age of 92. There is no need for me to remind my fellow members of the American Philosophical Association of his great contributions to philosophy. But I will take a moment to express something about what he meant to me personally. He was not just my teacher and my friend. He was my hero, a man who was quietly but passionately committed to truth, to clarity, to understanding everything under the sun—and to making himself understood. More than anybody else he has made me proud to be a philosopher, so I would like to dedicate my Presidential Address to his memory.

*Parfois je pense; et parfois, je suis.*  
Paul Valéry<sup>1[1]</sup>

Valéry's "*Variation sur Descartes*" excellently evokes the vanishing act that has haunted philosophy ever since Darwin overturned the Cartesian tradition. If *my body* is composed of nothing but a team of a few trillion robotic cells, mindlessly interacting to produce all the large-scale patterns that tradition would attribute to the non-mechanical workings of my mind, there seems to be nothing left over to be *me*. Lurking in Darwin's shadow there is a bugbear: the incredible Disappearing Self.<sup>2[2]</sup> One of Darwin's earliest critics saw what was coming and could scarcely contain his outrage:

In the theory with which we have to deal, Absolute Ignorance is the artificer; so that we may enunciate as the fundamental principle of the whole system, that, IN ORDER TO MAKE A PERFECT AND BEAUTIFUL MACHINE, IT IS NOT REQUISITE TO KNOW HOW TO MAKE IT. This proposition will be found, on careful examination, to express, in condensed form, the essential purport of the Theory, and to express in a few words all Mr. Darwin's meaning; who, by a strange inversion of reasoning, seems to think Absolute Ignorance fully qualified to take the place of Absolute Wisdom in all the achievements of creative skill.<sup>3[3]</sup>

---

<sup>1[1]</sup>Paul Valéry, *Cahiers*, ed. Judith Robinson, 2 vols. (Paris: Edition de la Pleiade 1973-74). vol. 2 (1974), p. 1388.

<sup>2[2]</sup>Dennett, D. 1984, *Elbow Room*, p13.

<sup>3[3]</sup>Robert Beverley MacKenzie, 1868, *The Darwinian Theory of the Transmutation of Species Examined* (published anonymously "By a Graduate of the University of Cambridge"). London: Nisbet & Co. Quoted in a review, *Athenaeum*, no 2102, February 8, p217.

This “strange inversion of reasoning” promises—or threatens—to dissolve the Cartesian *res cogitans* as the wellspring of creativity, and then where will we be? Nowhere, it seems. It *seems* that if creativity gets “reduced” to “mere mechanism” *we* will be shown not to exist at all. Or, we will exist, but we won’t be thinkers, we won’t manifest genuine “Wisdom in all the achievements of creative skill.” The individual as Author of works and deeds will be demoted: a person, it seems, is a barely salient nexus, a mere slub in the fabric of causation.

Whenever we zoom in on the act of creation, it seems we lose sight of it. The genius we thought we could see from a distance gets replaced at the last instant by stupid machinery, an echo of Darwin’s shocking substitution of Absolute Ignorance for Absolute Wisdom in the creation of the biosphere. Many people dislike Darwinism in their guts, and of all the ill-lit, murky reasons for antipathy to Darwinism, this one has always struck me as the deepest, but only in the sense of being the most entrenched, the least accessible to rational criticism. There are thoughtful people who scoff at Creationism, dismiss dualism out of hand, pledge allegiance to academic humanism—and then get quite squirrely when somebody proposes a Darwinian theory of creative intelligence. The very idea that all the works of human genius can be understood *in the end* to be mechanistically generated products of a cascade of generate-and-test algorithms arouses deep revulsion in many otherwise quite insightful, open-minded people.

Absolute Ignorance? Fie on anybody who would thus put “A” and “I” together! Serendipity is the wellspring of evolution, so it is fitting that an evolutionist such as I should adapt MacKenzie’s happy capitalization for a purpose he could hardly have imagined. His outraged scoffing at the powers of Absolute Ignorance has an uncannily similar echo more than a century later in the equally outraged scoffing at those who believe in what John Searle<sup>4[4]</sup> has called “strong AI,” the thesis that *real* intelligence can be made by artifice, that the difference between a mindless mechanism and a mindful one is a difference of design (or *program*—since whatever you can design in hardware you can implement in a virtual machine that has the same competence)<sup>5[5]</sup>.

---

<sup>4[4]</sup>Searle, John, 1980, "Minds, Brains and Programs," Behavioral and Brain Sciences, 3, pp.417-58.

<sup>5[5]</sup>This is obviously true of all competences of information-processing or control, but not of productive or transformative processes, such as lactation, which requires the transport and assembly of particular materials. Since Searle purports to distinguish the brain’s “control powers” from its “bottom-up causal powers” that “produce intentionality,” some have thought Searle imagines intentionality to be a special sort of substance secreted by the brain. Since he denies this, he owes us some other way to distinguish these mysterious causal powers from the control powers that software can implement and an explanation of why they are not implementable in a virtual machine.

Darwin's "strange inversion of reasoning" turns an ancient idea upside-down. The "top-down" perspective on creative intelligence supposes that it always takes a big, fancy, smart thing to create a lesser thing. No horseshoe has ever made a blacksmith; no pot has fashioned a potter. Hence we—and all the other fancy things we see around us—must have been created by something still fancier, something like us only more so. To many—perhaps most—people, this idea is *just obvious*. Consider this page from a creationist propaganda mailing:

1. Do you know of any building that didn't have a builder? YES/NO
2. Do you know of any painting that didn't have a painter? YES/NO
3. Do you know of any car that didn't have a maker? YES/NO

If you answered "YES" for any of the above give details: \_\_\_\_\_

But however strongly the idea appeals to common sense, Darwin shows us how it can be, in a word, false. Darwin shows us that a bottom-up theory of creation is, indeed, not only imaginable but empirically demonstrable. Absolute Ignorance *is* fully qualified to take the place of Absolute Wisdom in all the achievements of creative skill—all of them.

John Searle's Chinese Room thought experiment is a variation on the desperate joke of the creationists:

Do you know of any machine that can understand Chinese? YES/NO

If you answered "YES" give details! \_\_\_\_\_

While the creationists' rhetorical questions merely gesture towards the presumed embarrassments facing anybody who tries to "give details" of an instance of bottom-up creation, Searle's challenge offers a survey of possible avenues the believers in strong AI might take in their attempts to "give details" and purports to rebut them one and all. The believers in strong AI have been remarkably unmoved by Searle's attempts at refutation, and the comparison of Searle's position with creationism shows why. Biologists who cannot *yet* explain some particular puzzle about the non-miraculous path that led to one marvel of nature or another, who cannot *yet* "give details" to satisfy the particular critic, nevertheless have such a fine track record of success in giving the details, and such a stable and fecund background theory to use in generating and confirming new details, that they simply dismiss the rhetorical implication: "You'll never succeed!" They calmly acknowledge that they may need to develop a few new wrinkles before they can declare victory. Believers in strong AI are similarly content to concede that all AI models to date have been deficient in many respects, orders of magnitude too simple, many of them pursuing particular visions of AI that are simply mistaken. They go on to note that Searle isn't challenging particular details of the attempts to date; he purports to be offering an argument for the *in principle* impossibility of strong AI, a conclusion that he insists is meant to cover all *imaginable* complications of the underlying theoretical framework. They know that their underlying theoretical framework is nothing other than the straightforward extension, into the human brain and all its peripheral devices and interfaces, of the Darwinian program of mindless mechanism doing, in the end, all the work. If Darwinian mechanisms can explain the existence of a skylark, in all its glory, they can surely explain the existence of an ode to a nightingale, too.<sup>6[6]</sup> A poem is a wonderful thing, but not clearly more wonderful than a living, singing skylark.

Unsupportable antipathies often survive thanks to protective coloration: they blend into the background of legitimate objections to overstatements of the view under attack. Since the reach of Darwinian enthusiasm has always exceeded its grasp, there are always good criticisms of Darwinian excesses to hide amongst. Likewise, of course, for the excesses of the ideologues of AI. And so the battle rages, generating as much suspicion as insight. Darwinians who are sure that a properly nuanced, sophisticated Darwinism is proof against all the objections and misgivings—I am one such—should nevertheless recall the fate of the Freudian nags of the 50s and 60s, who insisted on seeing everything through the perspective of their hero's categories, only to discover that by the time you've

---

<sup>6[6]</sup>This perspective helps to explain the visceral appeal to many onlookers of the various *apparent* alternatives to Darwinian mechanism that have flourished over the years. The most prominent recently have been the appeal to "self-organization" "on the edge of chaos" (Stuart Kauffman, Per Bak, and others), and "dynamical systems theory" in both evolution and cognition (Esther Thelen, Walter Freeman, Timothy van Gelder, and others), and, of course, Stephen Jay Gould's insistence that evolution is not, as I have claimed (building on the work of theorists from Darwin to Fisher and Haldane to Williams and Maynard Smith), fundamentally an algorithmic process. After the smoke of battle clears, these ideas can be readily seen to be, at best, interesting complications of the basic Darwinian mechanisms, just as connectionist architectures and embodied cognition models are interesting complications of the basic ideas of AI. These controversies are, at best, constructive disagreements over how to "give the details", not challenges to the basic Darwinian vision.



attenuated your Freudianism to accommodate *everything*, it is Pickwickian Freudianism most of the way. Sometimes a cigar is just a cigar, and sometimes an idea is just an idea—not a meme—and sometimes a bit of mental machinery is not usefully interpreted as an adaptation dating back to our ancestral hunter-gatherer days or long before, even though it is, obviously, descended (with modifications) from some combination or other of such adaptations. We Darwinians will try to remind ourselves of this, hoping our doughty opponents will come to recognize that a Darwinian theory of creativity is not just a promising solution but the only solution in sight to a problem that is everybody's problem: *how can an arrangement of a hundred billion mindless neurons compose a creative mind, an I?*

William Poundstone has put the inescapable challenge succinctly in terms of “the old fantasy of a monkey typing *Hamlet* by accident.” He calculates that the chances of this happening are “1 in 50 multiplied by itself 150,000 times.”

In view of this, it may seem remarkable that anything as complex as a text of *Hamlet* exists. The observation that *Hamlet* was written by Shakespeare and not some random agency only transfers the problem. Shakespeare, like everything else in the world, must have arisen (ultimately) from a homogeneous early universe. Any way you look at it *Hamlet* is a product of that primeval chaos.<sup>7[7]</sup>

Where does all that design come from? What processes could conceivably yield such improbable “achievements of creative skill”? What Darwin saw is that design is always both valuable and costly. It does not fall like manna from heaven, but must be accumulated the hard way, by time-consuming, energy-consuming processes of mindless search through “primeval chaos”, automatically preserving happy accidents when they occur. This broadband process of Research and Development is breathtakingly inefficient, but—this is Darwin's great insight—the costly fruits of R and D can be thriftily conserved, copied, and re-used, they can be accumulated over time to yield “the achievements of creative skill.” “This principle of preservation,” Darwin says, “I have called, for the sake of brevity, Natural Selection.”<sup>8[8]</sup>

There is no requirement in Darwin's vision that these R and D processes run everywhere and always at the same tempo, with the same (in-)efficiency. Consider the unimaginably huge multi-dimensional space of all *possible* designed things—both natural and artificial. Every imaginable whale and unicorn, every automobile and spaceship and robot, every poem and mathematical proof and symphony finds its place somewhere in this Design Space. If we think of design work or R and D as a sort of *lifting* in Design Space<sup>9[9]</sup> then we can see that the gradualistic, frequently back-sliding, maximally inefficient basic search process can on important occasions yield new conditions that speed up the process, permitting faster, more effective local lifting.<sup>10[10]</sup> Call any such product of earlier R and D a *crane*, and distinguish it from what Darwinism says does not happen: *skyhooks*.<sup>11[11]</sup> Skyhooks, like manna from heaven, would be miracles, and if we posit a skyhook anywhere in our “explanation” of creativity, we have in fact conceded defeat—“Then a miracle occurs.”<sup>12[12]</sup>

What, then, is a mind? The Darwinian answer is straightforward. A mind is a crane, made of cranes, made of cranes, a mechanism of not quite unimaginable complexity that can clamber through Design Space at a giddy—but not miraculously giddy—pace, thanks to all the earlier R and D, from all sources, that it exploits. What is the anti-

<sup>7[7]</sup>William Poundstone, 1985, *The Recursive Universe: Cosmic Complexity and the Limits of Scientific Knowledge*, New York: Wm Morrow, p23.

<sup>8[8]</sup>Charles Darwin, *Origin of Species*, Ch. 4 summary (p.127 in facs. edition)

<sup>9[9]</sup>This tactic of mapping evolutionary processes and results onto space is a natural and oft-used metaphor, exploited in models of hill-climbing, and peaks in adaptive landscapes, to name the most obvious and popular applications. Its naturalness does not guarantee its soundness, of course, and may even mask its limitations, but since the basic mapping strategy has proven to be particularly useful in expressing *criticisms* of over-simple evolutionary ideas (e.g., Kauffman's “rugged landscape”, Eigen's “quasi-species”), it is not obviously biased in favor of simplistic visions of Darwinism.

<sup>10[10]</sup>John Maynard Smith and Eörs Szathmáry, 1995, *The Major Transitions in Evolution*, Oxford: Freeman, identify eight occasions (major transitions) when the evolutionary process became more efficient, creating cranes.

<sup>11[11]</sup>Daniel C. Dennett, *Darwin's Dangerous Idea*, 1995, “The Tools for R and D: Skyhooks or Cranes?” pp 73-80.

<sup>12[12]</sup>See the famous cartoon by Sydney Harris, in which the physicist's blackboard is covered with impressive formulae, except for this bracketed phrase in the middle, which leads the onlooker scientist to say “I think you should be more explicit here in step two.” (reprinted in Daniel Dennett, *Consciousness Explained*, 1991, p38).

Darwinian answer? It is perfectly expressed by one of the 20<sup>th</sup> century's great creative geniuses (though, like MacKenzie, he probably didn't mean by his words what I intend to mean by them).

*Je ne cherche pas; je trouve.*  
—Pablo Picasso

Picasso purports to be a genius indeed, someone who does not need to engage in the menial work of trial and error, generate-and-test, R and D; he claims to be able to *leap* to the summits of the peaks—the excellent designs—in the vast reaches of Design Space without having to guide his trajectory (he searches not) by sidelong testing at any waystations. As an inspired bit of bragging, this is *non pareil*, but I don't believe it for a minute. And anyone who has strolled through an exhibit of Picasso drawings (as I recently did in Valencia) looking at literally dozens of variations on a single theme, all signed—and sold—by the artist, will appreciate that whatever Picasso may have meant by his *bon mot*, he could not truly claim that he didn't engage in a time-consuming, energy-consuming exploration of neighborhoods in Design Space. At best he could claim that his own searches were so advanced, so efficient, that it didn't seem—to himself—to be design *work* at all. But then what did he have within him that made him such a great designer? A skyhook, or a superb collection of cranes?<sup>13[13]</sup>

We can now characterize a mutual suspicion between Darwinians and anti-Darwinians that distorts the empirical investigation of creativity. Darwinians suspect their opponents of hankering after a skyhook, a miraculous gift of genius whose powers have no decomposition into mechanical operations, however complex and informed by earlier processes of R and D. Anti-Darwinians suspect their opponents of hankering after an account of creative processes that so diminishes the Finder, the Author, the Creator, that it disappears, at best a mere temporary locus of mindless differential replication. We can make a little progress, I think, by building on Poundstone's example of the *creation of the creator of Hamlet*. Consider, then, a little thought experiment.

Suppose Dr. Frankenstein designs and constructs a monster, Spakesheare, that thereupon sits up and writes out a play, *Spamlet*. My question is not about the author of *Waverley* but about the author of *Spamlet*.

Who is the author of *Spamlet*?

First, let's take note of what I claim to be irrelevant in this thought experiment. I haven't said whether Spakesheare is a robot, constructed out of metal and silicon chips, or, like the original Frankenstein's monster, constructed out of human tissues—or cells, or proteins, or amino acids, or carbon atoms. As long as the design work and the construction were carried out by Dr. Frankenstein, it makes no difference to the example what the materials are. It might well turn out that the only way to build a robot small enough and fast enough and energy-efficient enough to sit on a stool and type out a play is to construct it from artificial cells filled with beautifully crafted motor proteins and other carbon-based nanorobots. That is an interesting technical and scientific question, but not of concern here. For exactly the same reason, if Spakesheare is a metal-and-silicon robot, it may be allowed to be larger than a galaxy, if that's what it takes to get the requisite complication into its program—and we'll just have to repeal the speed limit for light for the sake of our thought experiment. These technical constraints are commonly declared to be off-limits in these thought experiments, so so be it. If Dr. Frankenstein chooses to make his AI robot out of proteins and the like, that's his business. If his robot is cross-fertile with normal human beings and hence capable of creating what is arguably a new species by giving birth to a child, that is fascinating, but what we will be concerned with is Spakesheare's purported brainchild, *Spamlet*. Back to our question:

Who is the author of *Spamlet*?

In order to get a grip on this question, we have to look inside and see what happens in Spakesheare.<sup>14[14]</sup> At one extreme, we find inside a file (if Spakesheare is a robot with a computer memory) or a basically *memorized* version of *Spamlet*, all loaded and ready to run. In such an extreme case, Dr. Frankenstein is surely the author of *Spamlet*<sup>15[15]</sup>, using his intermediate creation, Spakesheare, as a mere storage-and-delivery device, a particularly fancy word processor. All the R and D work was done earlier, and copied to Spakesheare by one means or another.

<sup>13[13]</sup>I have been unable to discover the source of Picasso's claim, which is nicely balanced by a better known remark by a more down-to-earth creative genius, Thomas Edison: "Genius is one per cent. inspiration and ninety-nine per cent. perspiration." (in a newspaper interview. *Life* [1932], ch. 24, according to the *Oxford Dictionary of Quotations*.)

<sup>14[14]</sup>Yes, I intend the homage to an old favorite of mine, *What Happens in Hamlet*, by J. Dover Wilson (1951, Cambridge Univ. Press).

<sup>15[15]</sup>unless we find there is a Ms. Shelley who is the author of Dr. Frankenstein. . . . !

We can visualize this more clearly by imagining a sub-space of Design Space, which I call the Library of Babel, after Jorge Luis Borges' classic short story by that name.<sup>16[16]</sup> Borges invites us to imagine a warehouse filled with books that appears to its inhabitants to be infinite; they eventually decide that it is not, but it might as well be, for it seems that on its shelves--in no order, alas--lie all the *possible* books.

Suppose that each book is 500 pages long, and each page consists of 40 lines of 50 spaces, so there are two thousand character-spaces per page. Each space is either blank, or has a character printed on it, chosen from a set of 100 (the upper and lower case letters of English and other European languages, plus the blank and punctuation marks).<sup>17[17]</sup> Somewhere in the Library of Babel is a volume consisting entirely of blank pages, and another volume is all question marks, but the vast majority consist of typographical gibberish; no rules of spelling or grammar, to say nothing of sense, prohibit the inclusion of a volume. Five hundred pages times two thousand characters per page gives a million character-spaces per book, so there are  $100^{1,000,000}$  books in the Library of Babel. Since it is estimated<sup>18[18]</sup> that there are only  $100^{40}$  (give or take a few) *particles* (protons, neutrons and electrons) in the region of the universe we can observe, the Library of Babel is not remotely a physically possible object, but thanks to the strict rules with which Borges constructed it in his imagination, we can think about it clearly.

We need some terms for the quantities involved. The Library of Babel is not infinite, so the chance of finding anything interesting in it is not literally infinitesimal.<sup>19[19]</sup> These words exaggerate in a familiar way, but we should avoid them. Unfortunately, all the standard metaphors--astronomically large, a needle in a haystack, a drop in the ocean--fall comically short. No *actual* astronomical quantity (such as the number of elementary particles in the universe, or the time since the Big Bang measured in nanoseconds) is even visible against the backdrop of these huge-but-finite numbers. If a readable volume in the Library were as easy to find as a particular drop in the ocean, we'd be in business! Dropped at random into the Library, your chance of ever encountering a volume with so much as a grammatical sentence in it is so vanishingly small that we might do well to capitalize the term--*Vanishingly* small--and give it a mate, *Vastly*, short for Very-much-more-than-astronomically.<sup>20[20]</sup>

It is amusing to reflect on just how large this finite set of possible books is, compared with any actual library. Most of the books are pure gibberish, as noted, so consider the Vanishing subset of books composed entirely

<sup>16[16]</sup>Borges, J. L. 1962, *Labyrinths: Selected Stories and other Writings*, New York: New Directions. [La Biblioteca de Babel, 1941, in *El jardin de los senderos que se bifurcan*, published with another in *Ficciones*. 1956, Emece Editores, S. A., Buenos Aires.]

<sup>17[17]</sup>Borges chose slightly different figures: books 410 pages long, with 40 lines of 80 characters. The total number of characters per book is close enough to mine (1,1312,000 versus 1,000,000) to make no difference. I chose my rounder numbers for ease of handling. Borges chose a character set with only 25 members, which is enough for upper-case Spanish (with a blank, a comma and a period as the only punctuation), but not for English. I chose the more commodious 100 to make room without any doubt for the upper and lower case letters and punctuation of all the Roman alphabet languages.

<sup>18[18]</sup> Steven Hawking insists on putting it this way: "There are something like ten million million million million million million million million million million (1 with eighty zeroes after it) particles in the region of the universe that we can observe." *A Brief History of Time*, New York: Bantam, p.129. Michael Denton (*Evolution: A Theory in Crisis*, London: Burnett Books. 1985) provides the estimate of  $10^{70}$  atoms in the observable universe. Manfred Eigen (*Steps Towards Life*, Oxford University Press 1992, p.10) calculates the volume of the universe as  $10^{84}$  cubic centimeters.

<sup>19[19]</sup> The Library of Babel is finite, but curiously enough, it contains all the grammatical sentences of English within its walls. But that's an infinite set, and the library is finite! Still, any sentence of English, of whatever length, can be broken down into 500-page chunks, each of which is somewhere in the library! How is this possible? Some books may get used more than once. The most profligate case is the easiest to understand: since there are volumes which each contain a single character and are otherwise blank, repeated use of these one hundred volumes will create any text of any length. As Quine (*Quiddities: An Intermittently Philosophical Dictionary*, Cambridge, MA: Harvard Univ. Press. 1987) points out, in his informative and amusing essay, "Universal Library," if you avail yourself of this strategy of re-using volumes, and translate everything into the ASCII code your wordprocessor uses, you can store the whole Library of Babel in two extremely slender volumes, in one of which is printed a 0 and in the other of which appears a 1! (Quine also points out that Theodor Fechner, the psychologist, propounded the fantasy of the universal library long before Borges.)

<sup>20[20]</sup> Quine, *loc.cit.*, coins the term "hyperastronomic" for the same purpose. The previous two paragraphs are drawn, with minor changes, from *Darwin's Dangerous Idea*, pp108-9.

of English words, without a single misspelling. It is itself a Vast set, of course, and contained within it, but Vanishingly hard to find, is the Vast subset whose English words are lined up in grammatical sentences. A Vast but Vanishing subset of this subset in turn is the subset of books composed of English sentences that actually make sense. A Vast but Vanishing subset of these are about somebody named John, and a Vast but Vanishing subset of these are about the death of John F. Kennedy. A Vast but Vanishing subset of these are true . . . and a Vast but Vanishing subset of the possible true books about the death of JFK are written entirely in limericks. There are many orders of magnitude more possible true books in limerick form about the death of JFK than there are books in the Library of Congress.

Now we are ready to return to that needle-in-a-haystack, *Spamlet*, and consider how the trajectory to this particular place in the Library of Babel was traversed in actual history. If we find that the whole journey was already completed by the time Spakesheare's memory was constructed and filled with information, we know that Spakesheare played no role at all in the search. Working backwards, if we find that Spakesheare's only role was running the stored text through a spell-checker before using it to guide its typing motions, we will be unimpressed by claims of Spakeshearian authorship. This is a measurable, but Vanishing, part of the total R and D. There is a sizable galaxy of near-twin texts of *Spamlet*--roughly a hundred million different minor mutants have but a single uncorrected typo in them, and if we expand our horizon to include one typo per page, we have begun to enter the land of Vast numbers of variations on the theme. Working back a little further, once we graduate from typos to *thinkos*,<sup>21[21]</sup> those arguably mistaken, or sup-optimally chosen, words, we have begun to enter the land of serious authorship, as contrasted with mere copy-editing. The relative triviality of copy-editing, and yet its unignorable importance in shaping the final product gets well represented in terms of our metaphor of Design Space, where every little bit of lifting counts for something, and sometimes a little bit of lifting moves you onto a whole new trajectory. As usual, we may quote Ludwig Mies van der Rohe at this juncture: "God is in the details."

Now let's turn the knobs on our thought experiment, as Douglas Hofstadter has recommended<sup>22[22]</sup> and look at the other extreme, in which Dr. Frankenstein leaves most of the work to Spakesheare. The most realistic scenario would surely be that Spakesheare has been equipped by Dr. Frankenstein with a virtual past, a lifetime stock of pseudo-memories of experiences on which to draw while responding to its Frankenstein-installed obsessive desire to write a play. Among those pseudo-memories, we may suppose, are many evenings at the theater, or reading books, but also some unrequited loves, some shocking close calls, some shameful betrayals and the like. Now what happens? Perhaps some scrap of a "human interest" story on the network news will be the catalyst that spurs Spakesheare into a frenzy of generate-and-test, ransacking its memory for useful tidbits and themes, transforming--transposing, morphing--what it finds, jiggling the pieces into temporary, hopeful structures that compete for completion, most of them dismantled by the corrosive processes of criticism that nevertheless expose useful bits now and then, and so forth, and all of this multi-leveled search would be somewhat guided by multi-level, internally generated evaluations, including evaluation of the evaluation . . . of the evaluation functions as a response to evaluation of . . . the products of the ongoing searches.<sup>23[23]</sup>

Now if the amazing Dr. Frankenstein had actually anticipated all this activity down to its finest grain at the most turbulent and chaotic level, and had hand-designed Spakesheare's virtual past, and all its search machinery, to yield just this product, *Spamlet*, then Dr. Frankenstein would be, once again, the author of *Spamlet*, but also, in a word, God. Such Vast foreknowledge would be simply miraculous. Restoring a smidgen of realism to our fantasy, we can set the knobs at a rather less extreme position and assume that Dr. Frankenstein was unable to foresee all this in detail, but rather delegated to Spakesheare most of the hard work of completing the trajectory in Design Space to *one literary work or another*, something to be determined by later R and D occurring within Spakesheare itself. We have now arrived, by this simple turn of the knob, in the neighborhood of reality itself, for we already have actual examples of impressive artificial Authors that Vastly outstrip the foresight of their own creators. Nobody has yet created an artificial playwright worth serious attention, but an artificial chess player--IBM's Deep Blue--and an artificial composer--David Cope's EMI--have both achieved results that are, *in some respects*, equal to the best that human creative genius can muster.

Who beat Garry Kasparov, the reigning World Chess Champion? Not Murray Campbell or any of his IBM

<sup>21[21]</sup>For more on this concept, see my "From Typo to Thinko: When Evolution Graduated to Semantic Norms," forthcoming in the Fyssen conference volume on cultural evolution.

<sup>22[22]</sup>Douglas R. Hofstadter, 1981, "Reflections," (on Searle) in Hofstadter and Dennett, eds., *The Mind's I: Fantasies and Reflections on Self and Soul*, 1981, New York: Basic Books and Hassocks, Sussex: Harvester

<sup>23[23]</sup>Shakespeare himself was, of course, a tireless exploiter of the design work of others, and may well have been poking fun at his own reputation, quoting a critic, when he had Autolycus describe himself as "a snapper-up of unconsidered trifles" in *A Winter's Tale* (Act IV, scene iii). Thanks to Tony Marcel for drawing this passage to my attention.



team. Deep Blue beat Kasparov. Deep Blue designs better chess games than any of them can design. None of them can author a winning game against Kasparov. Deep Blue can. Yes, but. Yes, but. I am sure many of you are tempted to insist at this point that when Deep Blue beats Kasparov at chess, its brute force search methods are *entirely* unlike the exploratory processes that Kasparov uses when he conjures up his chess moves. But that is simply not so—or at least it is not so in the only way that could make a difference to the context of this debate about the universality of the Darwinian perspective on creativity. Kasparov’s brain is made of organic materials, and has an architecture importantly unlike that of Deep Blue, but it is still, so far as we know, a massively parallel search engine which has built up, over time, an outstanding array of heuristic pruning techniques that keep it from wasting time on unlikely branches. There is no doubt that the investment in R and D has a different profile in the two cases; Kasparov has methods of extracting good design principles from past games, so that he can recognize, and know enough to ignore, huge portions of the game space that Deep Blue must still patiently canvass *seriatim*. Kasparov’s “insight” dramatically changes the shape of the search he engages in, but it does not constitute “an *entirely* different” means of creation. Whenever Deep Blue’s exhaustive searches close off a *type* of avenue that it has some means of recognizing (a difficult, but not impossible task), it can re-use that R and D whenever it is appropriate, just as Kasparov does. Much of this analytical work has been done for Deep Blue by its designers, and given as an innate endowment, but Kasparov has likewise benefitted from hundreds of thousands of person-years of chess exploration transmitted to him by players, coaches and books. It is interesting in this regard to contemplate the suggestion recently made by Bobby Fischer, who proposes to restore the game of chess to its intended rational purity by requiring that the major pieces be *randomly* placed in the back row at the start of each game (random, but mirror image for black and white). This would instantly render the mountain of memorized openings almost entirely obsolete, for humans and machines alike, since only rarely would any of this lore come into play. One would be thrown back onto a reliance on fundamental principles; one would have to do more of the hard design work in real time—with the clock running. It is far from clear whether this change in rules would benefit human beings more than computers. It all depends on which type of chess player is relying most heavily on what is, in effect, rote memory—reliance *with minimal comprehension* on the R and D of earlier explorers.

The fact is that the search space for chess is too big for even Deep Blue to explore exhaustively in real time, so like Kasparov, it prunes its search trees by taking calculated risks, and like Kasparov, it often gets these risks pre-calculated. Both presumably do massive amounts of “brute force” computation on their very different architectures. After all, what do neurons know about chess? Any work *they* do must be brute force work of one sort or another.

It may seem that I am begging the question in favor of a computational, AI approach by describing the work done by Kasparov’s brain in this way, but the work has to be done somehow, and no *other* way of getting the work done has ever been articulated. It won’t do to say that Kasparov uses “insight” or “intuition” since that just means that Kasparov himself has no privileged access, no insight, into how the good results come to him. So, since nobody knows how Kasparov’s brain does it—least of all Kasparov—there is not yet any evidence at all to support the claim that Kasparov’s means are “entirely unlike” the means exploited by Deep Blue. One should remember this when tempted to insist that “of course” Kasparov’s methods are hugely different. What on earth could provoke one to go out on a limb like that? Wishful thinking? Fear?

But that’s just chess, you say, not art. Chess is *trivial* compared to art (now that the world champion chess player is a computer). This is where David Cope’s EMI comes into play.<sup>24[24]</sup> Cope set out to create a mere efficiency-enhancer, a composer’s aid to help him over the blockades of composition any creator confronts, a high-tech extension of the traditional search vehicles (the piano, staff paper, the tape recorder, etc.). As EMI grew in competence, it promoted itself into a whole composer, incorporating more and more of the generate-and-test process. When EMI is fed music by Bach, it responds by generating musical compositions in the style of Bach. When given Mozart, or Schubert, or Puccini, or Scott Joplin, it readily analyzes their styles and composes new music in their styles, better pastiches than Cope himself—or almost any human composer—can compose. When fed music by two composers, it can promptly compose pieces that eerily unite their styles, and when fed, all at once (with no clearing of the palate, you might say) all these styles at once, it proceeds to write music based on the totality of its musical experience. The compositions that result can then also be fed back into it, over and over, along with whatever other music comes along in MIDI format, and the result is EMI’s own “personal” musical style, a style that candidly reveals its debts to the masters, while being an unquestionably idiosyncratic integration of all this “experience.” EMI can now compose not just two-part inventions and art songs but whole symphonies—and has composed over a thousand, when last I heard. They are good enough to fool experts (composers and professors of music) and I can personally attest to the fact that an EMI-Puccini aria brought a lump to my throat—but then, I’m on a hair trigger when it comes to Puccini, and this was a good enough imitation to fool me. David Cope can no more claim to be the composer of EMI’s symphonies and motets and art songs than Murray Campbell can claim to have

<sup>24[24]</sup>For the details, see David Cope, ed., *Virtual Music* (forthcoming from MIT Press), including my commentary, “Collision Detection, Muselot, and Scribble: Some Reflections on Creativity.”



beaten Kasparov in chess.

To a Darwinian, this new element in the cascade of cranes is simply the latest in a long history, and we should recognize that the boundary between authors and their artifacts should be just as penetrable as all the other boundaries in the cascade. When Richard Dawkins notes<sup>25[25]</sup> that the beaver's dam is as much a part of the beaver phenotype—its *extended phenotype*—as its teeth and its fur, he sets the stage for the further observation that the boundaries of a human author are exactly as amenable to extension. In fact, of course, we've known this for centuries, and have carpentered various semi-stable conventions for dealing with the products of Rubens, of Rubens' *studio*, of Rubens' various students. Wherever there can be a helping hand, we can raise the question of just who is helping whom, what is creator and what is creation. How should we deal with such questions? To the extent that anti-Darwinians simply want us to preserve some tradition of authorship, to have some *rules of thumb* for determining who or what shall receive the honor (or blame) that attends authorship, their desires can be acknowledged and met, one way or another (which doesn't necessarily mean we should meet them). To the extent that this is not enough for the anti-Darwinians, to the extent that they want to hold out for authors as an objective, metaphysically grounded, "natural kind" (oh, the irony in those essentialist wolf-words in naturalist sheep's clothing!), they are looking for a skyhook.

The renunciation of skyhooks is, I think, the deepest and most important legacy of Darwin in philosophy, and it has a huge domain of influence, extending far beyond the skirmishes of evolutionary epistemology and evolutionary ethics. If we commit ourselves to Darwin's "strange inversion of reasoning," we turn our backs on compelling ideas that have been central to the philosophical tradition for centuries, not just Aristotle's essentialism and irreducible *telos*, but also Descartes's *res cogitans* as a causer outside the mechanistic world, to name the three that had been most irresistible until Darwin came along. The siren songs of these compelling traditions still move many philosophers who have not yet seen fit to execute the inversion, sad to say. Clinging to their pre-Darwinian assumptions, they create problems for themselves that will no doubt occupy many philosophers for years to come.<sup>26[26]</sup> The themes all converge when the topic is creativity and authorship, where the urge is to hunt for an "essence" of creativity, an "intrinsic" source of meaning and purpose, a locus of responsibility somehow insulated from the causal fabric in which it is embedded, so that within its boundaries it can generate, from its *own* genius<sup>27[27]</sup>, its *irreducible* genius, the meaningful words and deeds that distinguish us so sharply from mere mechanisms.

Plato called for us to carve nature at its joints, a wonderful biological image, and Darwin showed us that the salient boundaries in the biosphere are not the crisp set-theoretic boundaries of essentialism, but the emergent effects of historical processes. As one species turns into two, the narrow isthmus of intermediates disappears as time passes, leaving islands, concentrations sharing family resemblances, surrounded by empty space. As Darwin noted (in somewhat different terms), there are feedback processes that enhance separation, actively depopulating this middle ground. We might expect the same sort of effects in the sphere of human mind and culture, cultural habits or practices that favor the isolation of the processes of artistic creation in a single mind. "Are you the author of this?" "Is this all your own work?" The mere fact that these are familiar questions shows that there are cultural pressures encouraging people to *make* the favored answers come true. A small child, crayon in hand, huddled over her drawing, slaps away the helping hand of parent or sibling, because she wants this to be *her* drawing. She already appreciates the norm of pride of authorship, a culturally imbued bias built on the palimpsest of territoriality and biological ownership. The very idea of being an artist shapes her consideration of opportunities on offer, shapes her evaluation of features she discovers in herself. And this in turn will strongly influence the way she conducts her own searches through Design Space, in her largely unconscious emulation of Picasso's ideal, or, if she is of a contrarian spirit, defying it, like Marcel Duchamp:

<sup>25[25]</sup>Richard Dawkins, 1982, *The Extended Phenotype*, Oxford and San Francisco: Freeman.

<sup>26[26]</sup> Three examples: Jerry Fodor's series of flawed theories of psychosemantics; John Searle's inability to account for how "intrinsic intentionality" could evolve when it has no "control power" consequences visible to selective pressure; John McDowell's quest for a non-Darwinian alternative to what he calls "bald naturalism," a struggle to secure a variety of normativity that is not the mere *as-if* normativity he finds discernible in evolution. See Dennett, *Darwin's Dangerous Idea*, 1995, and "Granny versus Mother Nature -- No Contest," *Mind & Language*, **11** no.3, 1996, pp 263-269, and "Review of John Searle, *The Rediscovery of the Mind*" *Journal of Philosophy*, **60**, (4), 193-205, Apr. 1993, for my analyses of Fodor's and Searle's difficulties. My discussion of McDowell must be deferred to another occasion.

<sup>27[27]</sup>See "Do-It-Yourself Understanding," in *Brainchildren*, Cambridge MA: MIT Press, 1998, pp59-80, for my analysis of this theme in Fred Dretske's search for a privileged place where the understanding happens.

Cabanne: What determined your choice of readymades?

Duchamp: That depended on the object. In general, I had to beware of its “look.” It’s very difficult to choose an object, because, at the end of fifteen days, you begin to like it or to hate it. You have to approach something with an indifference, as if you had no aesthetic emotion. The choice of readymades is always based on visual indifference and, at the same time, on the total absence of good or bad taste. . . .<sup>28[28]</sup>

There is a persistent problem of imagination management in the debates surrounding this issue: people on both sides have a tendency to underestimate the resources of Darwinism, imagining simplistic alternatives that do not exhaust the space of possibilities. Darwinians are notoriously quick to find (or invent) differences in *genetic fitness* to go with every difference they observe, for instance. Meanwhile, anti-Darwinians, noting the huge distance between a beehive and the *St. Matthew Passion* as created objects, are apt to suppose that anybody who proposes to explain both creative processes with a single set of principles must be guilty of one reductionist fantasy or another: “Bach had a gene for writing baroque counterpoint just like the bees’ gene for forming wax hexagons” or “Bach was just a mindless trial-and-error mutator and selector of the musical memes that already flourished in his cultural environment.” Both of these alternatives are nonsense, of course, but pointing out their flaws does nothing to support the idea that (“therefore”) there must be irreducibly *non-Darwinian* principles at work in any account of Bach’s creativity. In place of this dimly imagined chasm with “Darwinian phenomena” on one side and “non-Darwinian phenomena” on the other side, we need to learn to see the space between bee and Bach as populated with all manner of mixed cases, differing from their nearest neighbors in barely perceptible ways, replacing the chasm with a traversable gradient of non-minds, protominds, hemi-demi-semi minds, magpie minds, copycat minds, aping minds, clever-pastiche minds, “path-finding” minds, “ground-breaking” minds, and eventually, genius minds. And the individual minds, of each caliber, will themselves be composed of different sorts of parts, including, surely, some special-purpose “modules” adapted to various new tricks and tasks, as well as a cascade of higher-order reflection devices, capable of generating ever more rarefied and delimited searches through pre-selected regions of the Vast space of possible designs.

It is important to recognize that genius is itself a product of natural selection and involves generate-and-test procedures all the way down. Once you have such a product, it is often no longer particularly perspicuous to view it solely as a cascade of generate-and-test processes. It often makes good sense to leap ahead on a *narrative* course, thinking of the agent as a self, with a variety of projects, goals, presuppositions, hopes, . . . . In short, it often makes good sense to adopt the intentional stance towards the whole complex product of evolutionary processes. This effectively brackets the largely unknown and unknowable mechanical microprocesses as well as the history that set them up, and puts them out of focus while highlighting the patterns of rational activity that those mechanical microprocesses track so closely. This tactic makes especially good sense to the creator himself or herself, who must learn not to be oppressed by the revelation that on close inspection, even on close *introspection*, a genius dissolves into a pack rat, which dissolves in turn into a collection of trial-and-error processes over which nobody has ultimate control.

Does this realization amount to a loss—an elimination—of selfhood, of genius, of creativity? Those who are closest to the issue—the artistic and scientific geniuses who have reflected on it—often confront this discovery with equanimity. Mozart is reputed to have said of his best musical ideas: “Whence and how do they come? I don’t know and I have nothing to do with it.”<sup>29[29]</sup> The painter Philip Guston is equally unperturbed by this evaporation of visible self when the creative juices start flowing:

When I first come into the studio to work, there is this noisy crowd which follows me there; it includes all of the important painters in history, all of my contemporaries, all the art critics, etc. As I become involved in the work, one by one, they all leave. If I’m lucky, every one of them will disappear. If I’m really lucky, I will too.<sup>30[30]</sup>

<sup>28[28]</sup> *Dialogues with Marcel Duchamp*, Pierre Cabanne (transl Ron Padgett), New York, Viking Press, 1971, p.48. Thanks to Nicholas Humphrey and Victoria McGeer for ideas expressed in the previous paragraph.

<sup>29[29]</sup> In an oft-quoted but possibly spurious passage—see *Darwin’s Dangerous Idea*, p346-7.

<sup>30[30]</sup> I have been unable to locate the source of Guston’s quote, but I have found much the same remark attributed to the composer, John Cage, a close friend and contemporary of Guston’s, who [is said to have] said this about painting: When you are working, everybody is in your studio—the past, your friends, the art world, and above all, your own ideas—all are there. But as you continue painting, they start leaving, one by one, and you are left completely alone. Then, if you are lucky, even you leave. Like all other creators, Guston and I like to re-use what we find, adding a few touches from time to time.

In closing, I would like to acknowledge a few of my co-authors:

Anonymous  
Jorge Luis Borges  
David Cope  
Charles Darwin  
Richard Dawkins  
Susan Dennett  
René Descartes  
Marcel Duchamp  
Thomas Edison  
Bobby Fischer  
Philip Guston  
Douglas Hofstadter  
Nicholas Humphrey  
Robert MacKenzie  
Tony Marcel  
Victoria McGeer  
Ludwig Mies van der Rohe  
Pablo Picasso  
William Poundstone  
John Searle  
William Shakespeare  
Mary Shelley  
Paul Valéry

---

## Postmodernism and Truth. FINAL DRAFT August 13, 1998

<http://ase.tufts.edu/cogstud/papers/postmod.tru.htm>

FINAL DRAFT

for World Congress of Philosophy

August 13, 1998

Daniel C. Dennett

### Postmodernism and Truth<sup>(1)</sup>

Here is a story you probably haven't heard, about how a team of American researchers inadvertently introduced a virus into a third world country they were studying. They were experts in their field, and they had the best intentions; they thought they were helping the people they were studying, but in fact they had never really seriously considered whether what they were doing might have ill effects. It had not occurred to them that a side-effect of their research might be damaging to the fragile ecology of the country they were studying. The virus they introduced had some dire effects indeed: it raised infant mortality rates, led to a general decline in the health and wellbeing of women and children, and, perhaps worst of all, indirectly undermined the only effective political force for democracy in the country, strengthening the hand of the traditional despot who ruled the nation. These American researchers had something to answer for, surely, but when confronted with the devastation they had wrought, their response was frustrating, to say the least: they still thought that what they were doing was, all things considered, in the interests of the people, and declared that the standards by which this so-called devastation was being measured were simply not appropriate. Their critics, they contended, were trying to impose "Western" standards in a cultural environment that had no use for such standards. In this strange defense they were warmly supported by the country's leaders--not

surprisingly--and little was heard--not surprisingly--from those who *might* have been said, by Western standards, to have suffered as a result of their activities.

These researchers were not biologists intent on introducing new strains of rice, nor were they agri-business chemists testing new pesticides, or doctors trying out vaccines that couldn't legally be tested in the U.S.A. They were postmodernist science critics and other multiculturalists who were arguing, in the course of their professional researches on the culture and traditional "science" of this country, that Western science was just one among many equally valid narratives, not to be "privileged" in its competition with native traditions which other researchers--biologists, chemists, doctors and others--were eager to supplant. The virus they introduced was not a macromolecule but a meme (a replicating idea): the idea that science was a "colonial" imposition, not a worthy substitute for the practices and beliefs that had carried the third-world country to its current condition. And the reason you have not heard of this particular incident is that I made it up, to dramatize the issue and to try to unsettle what seems to be current orthodoxy among the *literati* about such matters. But it is inspired by real incidents--that is to say, true reports. Events of just this sort have occurred in India and elsewhere, reported, movingly, by a number of writers, among them:

Meera Nanda, "The Epistemic Charity of the Social Constructivist Critics of Science and Why the Third World Should Refuse the Offer," in N. Koertge, ed., *A House Built on Sand: Exposing Postmodernist Myths about Science*, Oxford University Press, 1998, pp286-311

Reza Afshari, "An Essay on Islamic Cultural Relativism in the Discourse of Human Rights," in *Human Rights Quarterly*, **16**, 1994, pp.235-76.

Susan Okin, "Is Multiculturalism Bad for Women?" *Boston Review*, October/November, 1997, pp 25-28.

Pervez Hoodbhoy, *Islam and Science: Religious Orthodoxy and the Battle for Rationality*, London and New Jersey, Zed Books Ltd. 1991.

My little fable is also inspired by a wonderful remark of E. O. Wilson, in *Atlantic Monthly* a few months ago: "Scientists, being held responsible for what they say, have not found postmodernism useful." Actually, of course, we are all held responsible for what we say. The laws of libel and slander, for instance, exempt none of us, but most of us--including scientists in many or even most fields--do not typically make assertions that, independently of libel and slander considerations, might bring harm to others, even indirectly. A handy measure of this fact is the evident ridiculousness we discover in the idea of malpractice insurance for . . . literary critics, philosophers, mathematicians, historians, cosmologists. What on earth could a mathematician or literary critic do, in the course of executing her profession duties, that might need the security blanket of malpractice insurance? She might inadvertently trip a student in the corridor, or drop a book on somebody's head, but aside from such *outré* side-effects, our activities are paradigmatically innocuous. One would think. But in those fields where the stakes are higher--and more direct--there is a longstanding tradition of being especially cautious, and of taking

particular responsibility for ensuring that no harm results (as explicitly honored in the Hippocratic Oath). Engineers, knowing that thousands of people's safety may depend on the bridge they design, engage in focussed exercises with specified constraints designed to determine that, according to all current knowledge, their designs are safe and sound. Even economists--often derided for the risks they take with *other* people's livelihoods--when they find themselves in positions to endorse specific economic measures considered by government bodies or by their private clients, are known to attempt to put a salutary strain on their underlying assumptions, just to be safe. They are used to asking themselves, and to being expected to ask themselves: "What if I'm wrong?" We others seldom ask ourselves this question, since we have spent our student and professional lives working on topics that are, according both to tradition and common sense, incapable of affecting any lives in ways worth worrying about. If my topic is whether or not Vlastos had the best interpretation of Plato's *Parmenides* or how the wool trade affected imagery in Tudor poetry, or what the best version of string theory says about time, or how to recast proofs in topology in some new formalism, if I am wrong, dead wrong, in what I say, the only damage I am likely to do is to my own scholarly reputation. But when we aspire to have a greater impact on the "real" (as opposed to "academic") world-- and many philosophers do aspire to this today-- we need to adopt the attitudes and habits of these more applied disciplines. We need to hold ourselves responsible for what we say, recognizing that our words, if believed, can have profound effects for good or ill.

When I was a young untenured professor of philosophy, I once received a visit from a colleague from the Comparative Literature Department, an eminent and fashionable literary theorist, who wanted some help from me. I was flattered to be asked, and did my best to oblige, but the drift of his questions about various philosophical topics was strangely perplexing to me. For quite a while we were getting nowhere, until finally he managed to make clear to me what he had come for. He wanted "an epistemology," he said. *An* epistemology. Every self-respecting literary theorist had to sport an epistemology that season, it seems, and without one he felt naked, so he had come to me for an epistemology to wear--it was the very next fashion, he was sure, and he wanted the *dernier cri* in epistemologies. It didn't matter to him that it be sound, or defensible, or (as one might as well say) *true*; it just had to be new and different and stylish. Accessorize, my good fellow, or be overlooked at the party.

At that moment I perceived a gulf between us that I had only dimly seen before. It struck me at first as simply the gulf between being serious and being frivolous. But that initial surge of self-righteousness on my part was, in fact, a naive reaction. My sense of outrage, my sense that my time had been wasted by this man's bizarre project, was in its own way as unsophisticated as the reaction of the first-time theater-goer who leaps on the stage to protect the heroine from the villain. "Don't you understand?" we ask incredulously. "It's *make believe*. It's *art*. It isn't *supposed* to be taken literally!" Put in that context, perhaps this man's quest was not so disreputable after all. I would not have been offended, would I, if a colleague in the Drama Department had come by and asked if he could borrow a few yards of my books to put on the shelves of the set for his production of Tom Stoppard's play, *Jumpers*. What if anything would be wrong in outfitting this fellow with a snazzy set of outrageous epistemological doctrines with which he could titillate or confound his colleagues?

What would be wrong would be that since this man didn't acknowledge the gulf, didn't even recognize that it existed, my acquiescence in his shopping spree would have contributed to the debasement of a precious commodity, the erosion of a valuable distinction. Many people,



including both onlookers and participants, don't see this gulf, or actively deny its existence, and therein lies the problem. The sad fact is that in some intellectual circles, inhabited by some of our more advanced thinkers in the arts and humanities, this attitude passes as a sophisticated appreciation of the futility of proof and the relativity of all knowledge claims. In fact this opinion, far from being sophisticated, is the height of sheltered naiveté, made possible only by flatfooted ignorance of the proven methods of scientific truth-seeking and their power. Like many another naif, these thinkers, reflecting on the manifest inability of *their* methods of truth-seeking to achieve stable and valuable results, innocently generalize from their own cases and conclude that nobody *else* knows how to discover the truth either.

Among those who contribute to this problem, I am sorry to say, is, my good friend Dick Rorty. Richard Rorty and I have been constructively disagreeing with each other for over a quarter of a century now. Each of us has taught the other a great deal, I believe, in the reciprocal process of chipping away at our residual points of disagreement. I can't name a living philosopher from whom I have learned more. Rorty has opened up the horizons of contemporary philosophy, shrewdly showing us philosophers many things about how our own projects have grown out of the philosophical projects of the distant and recent past, while boldly describing and prescribing future paths for us to take. But there is one point over which he and I do not agree at all--not yet--and that concerns his attempt over the years to show that philosophers' debates about Truth and Reality really do erase the gulf, really do license a slide into some form of relativism. In the end, Rorty tells us, it is all just "conversations," and there are only political or historical or aesthetic grounds for taking one role or another in an ongoing conversation.

Rorty has often tried to enlist me in his campaign, declaring that he could find in my own work one explosive insight or another that would help him with his project of destroying the illusory edifice of objectivity. One of his favorite passages is the one with which I ended my book *Consciousness Explained* (1991):

It's just a war of metaphors, you say--but metaphors are not "just" metaphors; metaphors are the tools of thought. No one can think about consciousness without them, so it is important to equip yourself with the best set of tools available. Look what we have built with our tools. Could you have imagined it without them? [p.455]

"I wish," Rorty says, "he had taken one step further, and had added that such tools are all that inquiry can ever provide, because inquiry is never 'pure' in the sense of [Bernard] Williams' 'project of pure inquiry.' It is always a matter of getting us something we want." ("Holism, Intrinsicity, Transcendence," in Dahlbom, ed., *Dennett and his Critics*. 1993.) But I would never take that step, for although metaphors are indeed irreplaceable tools of thought, they are not the only such tools. Microscopes and mathematics and MRI scanners are among the others. Yes, any inquiry is a matter of getting us something we want: the truth about something that matters to us, if all goes as it should.

When philosophers argue about truth, they are arguing about how not to inflate the truth about truth into the Truth about Truth, some absolutistic doctrine that makes indefensible demands on

our systems of thought. It is in this regard similar to debates about, say, the reality of time, or the reality of the past. There are some deep, sophisticated, worthy philosophical investigations into whether, properly speaking, the past is real. Opinion is divided, but you entirely misunderstand the point of these disagreements if you suppose that they undercut claims such as the following:

Life first emerged on this planet more than three thousand million years ago.

The Holocaust happened during World War II.

Jack Ruby shot and killed Lee Harvey Oswald at 11:21 am, Dallas time, November 24, 1963.

These are truths about events that really happened. Their denials are falsehoods. No sane philosopher has ever thought otherwise, though in the heat of battle, they have sometimes made claims that could be so interpreted.

Richard Rorty deserves his large and enthralled readership in the arts and humanities, and in the "humanistic" social sciences, but when his readers enthusiastically interpret him as encouraging their postmodernist skepticism about truth, they trundle down paths he himself has refrained from traveling. When I press him on these points, he concedes that there is indeed a useful concept of truth that survives intact after all the corrosive philosophical objections have been duly entered. This serviceable, modest concept of truth, Rorty acknowledges, has its uses: when we want to compare two maps of the countryside for reliability, for instance, or when the issue is whether the accused did or did not commit the crime as charged.

Even Richard Rorty, then, acknowledges the gap, and the importance of the gap, between appearance and reality, between those theatrical exercises that may entertain us without pretence of truth-telling, and those that aim for, and often hit, the truth. He calls it a "vegetarian" concept of truth. Very well, then, let's all be vegetarians about the truth. Scientists never wanted to go the whole hog anyway.

So now, let's ask about the sources or foundations of this mild, uncontroversial, vegetarian concept of truth.

Right now, as I speak, billions of organisms on this planet are engaged in a game of hide and seek. It is not just a game for them. It is a matter of life and death. *Getting it right*, not making mistakes, has been of paramount importance to every living thing on this planet for more than three billion years, and so these organisms have evolved thousands of different ways of finding out about the world they live in, discriminating friends from foes, meals from mates, and ignoring the rest for the most part. It matters to them that they not be misinformed about these matters--indeed nothing matters more--but they don't, as a rule, appreciate this. They are the beneficiaries of equipment exquisitely designed to get what matters right but when their

equipment malfunctions and gets matters wrong, they have no resources, as a rule, for noticing this, let alone deploring it. They soldier on, unwittingly. The difference between how things seem and how things really are is just as fatal a gap for them as it can be for us, but they are largely oblivious to it. The *recognition* of the difference between appearance and reality is a human discovery. A few other species--some primates, some cetaceans, maybe even some birds--shows signs of appreciating the phenomenon of "false belief"--*getting it wrong*. They exhibit sensitivity to the errors of others, and perhaps even some sensitivity to their own errors as errors, but they lack the capacity for the reflection required to *dwell* on this possibility, and so they cannot use this sensitivity in the deliberate design of repairs or improvements of their own seeking gear or hiding gear. That sort of bridging of the gap between appearance and reality is a wrinkle that we human beings alone have mastered.

We are the species that discovered doubt. Is there enough food laid by for winter? Have I miscalculated? Is my mate cheating on me? Should we have moved south? Is it safe to enter this cave? Other creatures are often visibly agitated by their own uncertainties about just such questions, but because they cannot actually *ask themselves these questions*, they cannot articulate their predicaments for themselves or take steps to improve their grip on the truth. They are stuck in a world of appearances, making the best they can of how things seem and seldom if ever worrying about whether how things seem is how they truly are.

We alone can be wracked with doubt, and we alone have been provoked by that epistemic itch to seek a remedy: better truth-seeking methods. Wanting to keep better track of our food supplies, our territories, our families, our enemies, we discovered the benefits of talking it over with others, asking questions, passing on lore. We invented culture. Then we invented measuring, and arithmetic, and maps, and writing. These communicative and recording innovations come with a built-in ideal: truth. The point of asking questions is to find *true* answers; the point of measuring is to measure *accurately*; the point of making maps is to *find your way* to your destination. There may be an Island of the Colour-blind (allowing Oliver Sacks his usual large dose of poetic license), but no Island of the People Who Do Not Recognize Their Own Children. The Land of the Liars could exist only in philosophers' puzzles; there are no traditions of False Calendar Systems for mis-recording the passage of time. In short, the goal of truth goes without saying, in every human culture.

We human beings use our communicative skills not just for truth-telling, but also for promise-making, threatening, bargaining, story-telling, entertaining, mystifying, inducing hypnotic trances, and just plain kidding around, but prince of these activities is truth-telling, and for this activity we have invented ever better tools. Alongside our tools for agriculture, building, warfare, and transportation, we have created a technology of truth: science. Try to draw a straight line, or a circle, "freehand." Unless you have considerable artistic talent, the result will not be impressive. With a straight edge and a compass, on the other hand, you can practically eliminate the sources of human variability and get a nice clean, objective result, the same every time.

Is the line really straight? How straight is it? In response to these questions, we develop ever finer tests, and then tests of the accuracy of those tests, and so forth, bootstrapping our way to ever greater accuracy and objectivity. Scientists are just as vulnerable to wishful thinking, just as likely to be tempted by base motives, just as venal and gullible and forgetful as the rest of humankind. Scientists don't consider themselves to be saints; they don't even pretend to be

priests (who according to tradition are supposed to do a better job than the rest of us at fighting off human temptation and frailty). Scientists take themselves to be just as weak and fallible as anybody else, but recognizing those very sources of error in themselves and in the groups to which they belong, they have devised elaborate systems to tie their own hands, forcibly preventing their frailties and prejudices from infecting their results.

It is not just the implements, the physical tools of the trade, that are designed to be resistant to human error. The organization of methods is also under severe selection pressure for improved reliability and objectivity. The classic example is the double blind experiment, in which, for instance, neither the human subjects nor the experimenters themselves are permitted to know which subjects get the test drug and which the placebo, so that nobody's subliminal hankerings and hunches can influence the perception of the results. The statistical design of both individual experiments and suites of experiments, is then embedded in the larger practice of routine attempts at replication by independent investigators, which is further embedded in a tradition--flawed, but recognized--of publication of both positive and negative results.

What inspires faith in arithmetic is the fact that hundreds of scribblers, working independently on the same problem, will all arrive at the same answer (except for those negligible few whose errors can be found and identified to the mutual satisfaction of all). This unrivalled objectivity is also found in geometry and the other branches of mathematics, which since antiquity have been the very model of certain knowledge set against the world of flux and controversy. In Plato's early dialogue, the *Meno*, Socrates and the slave boy work out together a special case of the Pythagorean theorem. Plato's example expresses the frank recognition of a standard of truth to be aspired to by all truth-seekers, a standard that has not only never been seriously challenged, but that has been tacitly accepted--indeed heavily relied upon, even in matters of life and death--by the most vigorous opponents of science. (Or do you know a church that keeps track of its flock, and their donations, without benefit of arithmetic?)

Yes, but science almost never looks as uncontroversial, as cut-and-dried, as arithmetic. Indeed rival scientific factions often engage in propaganda battles as ferocious as anything to be found in politics, or even in religious conflict. The fury with which the defenders of scientific orthodoxy often defend their doctrines against the heretics is probably unmatched in other arenas of human rhetorical combat. These competitions for allegiance--and, of course, funding--are designed to capture attention, and being well-designed, they typically succeed. This has the side effect that the warfare on the cutting edge of any science draws attention away from the huge uncontested background, the dull metal heft of the axe that gives the cutting edge its power. What goes without saying, during these heated disagreements, is an organized, encyclopedic collection of agreed-upon, humdrum scientific fact.

Robert Proctor usefully draws our attention to a distinction between neutrality and objectivity.<sup>(2)</sup> Geologists, he notes, know a lot more about oil-bearing shales than about other rocks--for the obvious economic and political reasons--but they do *know* objectively about oil bearing shales. And much of what they learn about oil-bearing shales can be generalized to other, less favored rocks. We want science to be objective; we should not want science to be neutral. Biologists know a lot more about the fruit-fly, *Drosophila*, than they do about other insects--not because you can get rich off fruit flies, but because you can get knowledge out of fruit flies easier than you can get it out of most other species. Biologists also know a lot more about mosquitoes than about other insects, and here it is because mosquitoes are more harmful to people than other

species that might be much easier to study. Many are the reasons for concentrating attention in science, and they all conspire to making the paths of investigation far from neutral; they do not, in general, make those paths any less objective. Sometimes, to be sure, one bias or another leads to a violation of the canons of scientific method. Studying the pattern of a disease in men, for instance, while neglecting to gather the data on the same disease in women, is not just not neutral; it is bad science, as indefensible in scientific terms as it is in political terms.

It is true that past scientific orthodoxies have themselves inspired policies that hindsight reveals to be seriously flawed. One can sympathize, for instance, with Ashis Nandy, editor of the passionately anti-scientific anthology, *Science, Hegemony and Violence: A Requiem for Modernity*, Delhi: Oxford Univ. Press, 1988. Having lived through Atoms for Peace, and the Green Revolution, to name two of the most ballyhooed scientific juggernauts that have seriously disrupted third world societies, he sees how "the adaptation in India of decades-old western technologies are advertised and purchased as great leaps forward in science, even when such adaptations turn entire disciplines or areas of knowledge into mere intellectual machines for the adaptation, replication and testing of shop-worn western models which have often been given up in the west itself as too dangerous or as ecologically non-viable." (p8) But we should recognize this as a political misuse of science, not as a fundamental flaw in science itself.

The methods of science aren't foolproof, but they are indefinitely perfectible. Just as important: there is a tradition of criticism that enforces improvement whenever and wherever flaws are discovered. The methods of science, like everything else under the sun, are themselves objects of scientific scrutiny, as *method* becomes *methodology*, the analysis of methods. Methodology in turn falls under the gaze of *epistemology*, the investigation of investigation itself--nothing is off limits to scientific questioning. The irony is that these fruits of scientific reflection, showing us the ineliminable smudges of imperfection, are sometimes used by those who are suspicious of science as their grounds for denying it a privileged status in the truth-seeking department--as if the institutions and practices they see competing with it were no worse off in these regards. But where are the examples of religious orthodoxy being simply abandoned in the face of irresistible evidence? Again and again in science, yesterday's heresies have become today's new orthodoxies. No religion exhibits that pattern in its history.

1. Portions of this paper are derived from "Faith in the Truth," my Amnesty Lecture, Oxford, February 17, 1997

2. *Value-Free Science?*, Harvard Univ. Press, 1991

### Intentionality (1987)

<http://ase.tufts.edu/cogstud/papers/intentio.htm>

(with John Haugeland), in R. L. Gregory, ed., *The Oxford Companion to the Mind*, Oxford University Press 1987; reprinted in *Actes du 3ème Colloque International Cognition et Connaissance: Où va la science cognitive?* Toulouse: CNRS/Université Paul Sabatier 1988; reprinted in K. Lehrer and E. Sosa, eds., *The Opened Curtain: A U.S.-Soviet Philosophy Summit*, Westview Press, 1991, Chapter 3.

#### Intentionality

Daniel C. Dennett and John Haugeland

Intentionality is *aboutness*. Some things are about other things: a belief can be about icebergs, but an iceberg is not about anything; an idea can be about the number 7, but the number 7 is not about anything; a book or a film can be about Paris, but Paris is not about anything. Philosophers have long been concerned with the analysis of the phenomenon of intentionality, which has seemed to many to be a fundamental feature of mental states and events.



The term was coined by the Scholastics in the Middle Ages, and derives from the Latin verb *intendo*, meaning to point (at) or aim (at) or extend (toward). Phenomena with intentionality point outside themselves, in effect, to something else: whatever they are *of* or *about*. The term was revived in the 19th century by the philosopher and psychologist Franz Brentano, one of the most important predecessors of the school of Phenomenology. Brentano claimed that intentionality is the defining distinction between the mental and the physical; all and only mental phenomena exhibit intentionality. Since intentionality is, he claimed, an irreducible feature of mental phenomena, and since no physical phenomena could exhibit it, mental phenomena could not be a species of physical phenomena. This claim, often called the Brentano Thesis, or Brentano's Irreducibility Thesis, has often been cited to support the view that the mind cannot be the brain, but this is by no means generally accepted today.

There has been a second revival of the term in the last twenty years by English and American philosophers in the analytic tradition. In response to seminal work by R. Chisholm and W. V. O. Quine, a vigorous attempt has been made to develop an account of intentionality in harmony with the canons of modern logic and semantics. Since the Phenomenological tradition, mainly on the Continent, has continued to exploit the concept of intentionality along rather different lines, the problem of intentionality is one of the best points of convergent concern in these two largely separate--and often antagonistic--research traditions.

In spite of the attention currently devoted to the concept, there is a striking lack of received wisdom about the proper analysis of intentionality. What agreement there is concerns the nature of the problems raised. If we make an initial rough catalogue of the things that can be about things, it will include a great variety of *mental* states, and events (ideas, beliefs, desires, thoughts, hopes, fears, perceptions, dreams, hallucinations . . . ) but also various *linguistic* items (sentences, questions, poems, headlines, instructions . . . ) and perhaps other sorts of *representations* as well (pictures, charts, films, symphonic tone poems, computer programs, . . . ). Many have thought that these linguistic and nonlinguistic representations are only derivatively about anything. They depend for their intentionality on their being the creations and tools of creatures with minds, and particular representations derive their specific aboutness from the specific aboutness of the ideas, beliefs, and intentions of their creators. What makes the particular sequence of ink marks on the next line

Napoleon was exiled to Elba.

*about Napoleon* is its role as a sentence in a language used by people who *know* about Napoleon, or at least have *beliefs* about Napoleon, and *wish* to communicate about Napoleon. What *representations* mean depends on what *people* mean by using them. This suggests that the primary or underived intentionality of mental states and events is a very special feature indeed--the source of all meaning in the world. On this view sentences, pictures, diagrams, and the like are in effect prosthetic extensions of the minds of those who use them, having no intrinsic meaning but only the meaning they derive from their utilization.

Identifying intentionality with aboutness nicely locates the concept, but hardly clarifies it, for the ordinary word "about" is perplexing on its own. A belief can be about Paris, but a belief can also apparently be about phlogiston--and there is no phlogiston for it to be about. This curious fact, the possible non-existence of the *object* of an intentional item, may seem to be an idle puzzle, but in fact it has proven extraordinarily resistant to either solution or dismissal. Brentano called this the *intentional inexistence* of the intentional objects of mental states and events, and it has many manifestations. I cannot want without wanting something, but what I want need not exist for me to want it. It can be true that I now want a two-ton diamond even if there is not now and never has been or will be such a thing. People have believed in Poseidon, and children often believe in Santa Claus, and while in one sense we can say these believers all believe in nothing, their beliefs are quite different states of mind--they have different *objects*--and both are to be

distinguished from the state of mind of the skeptic or agnostic who can be said in quite a different sense to believe in nothing.

It might seem that there is a simple and straightforward solution to these problems: while Poseidon does not exist, an idea of Poseidon surely exists in each of the believers' minds, and this idea is the object of their belief. This will not do, however, for it is one thing to believe in the existence of the *idea* of Poseidon--we can all believe in the existence of that mental item--and quite another to believe in Poseidon. Similarly, what I might want could hardly be the idea in my mind of a twoton diamond, for that is something I already have. Moreover, when, as normally happens, the object of an intentional state does exist--e.g., when I believe that London is crowded-- that very object, London, the city itself and not my idea of it in my mind, is what my belief is about.

The relation, then, between a state of mind--or for that matter a sentence or picture--and its intentional object or objects is a peculiar relation in three ways.

First, for ordinary relations, like *x is sitting on y* or *x is employed by y*, *x* and *y* are identifiable entities quite apart from whether they happen to be thus related to each other. The thing which is *x* would be the same *x* whether or not it were sitting on *y* or employed by *y*. But the same is not true of intentional "relations". One and the same belief cannot at one moment be about a frog (that it is green, say) and at another moment be about a house (that it is green). The latter is a different belief. What a belief is supposed to be about is crucial to which belief it is.

Second, for ordinary relations, each of the things related must exist (or have existed), but as we have seen, intentional "relations" can be to non-existents.

Third, ordinary relations obtain between things regardless of how they might be specified. If I am sitting next to Jones and Jones is the Mad Strangler, then it follows that I am sitting next to the Mad Strangler, whatever anybody may think. But if I believe that Jones is harmless, or hope that he will marry my sister, it does not at all follow that I believe that the Mad Strangler is harmless or hope that the Mad Strangler will marry my sister. Even if one is tempted to object that in this case in one sense I do hope the Mad Strangler will marry her, there is clearly another sense in which I might hope this and don't.

For these reasons the normal logic of relations cannot accommodate the presumed relation between an intentional state and its intentional object or objects, but it has also not proven comfortable for theorists to deny on these grounds that there are such things as intentional relations--to hold that mental states, for instance, are only apparently relational. This, then, is the unsolved problem of intentionality.

Faced with this problem, the Anglo-American tradition, characteristically, has tended to favor a tactical retreat, to a logical analysis of the language we use to talk about intentional states, events and other items. This move, from the direct analysis of the phenomena to the analysis of our ways of talking about the phenomena, has been aptly called "semantic ascent" by Quine, and its immediate advantages are twofold. First, we set aside epistemological and metaphysical distractions such as: "How can we ever know another person's mental state anyway?" and "Are mental states a variety of physical state, or are they somehow immaterial or spiritual?". The things people say about mental states are in any event out in the public world where we can get at them and study them directly. Second, switching to language puts at our disposal a number of sophisticated techniques and theories developed by philosophers, logicians, and linguists. Semantic ascent is not guaranteed to solve any problems, of course, but it may permit them to be reformulated in ways more accessible to ultimate solution.

In its new guise, the problem of intentionality concerns the semantics of the so-called intentional idioms-- ". . . believes that p", ". . . desires that q", ". . . dreams that r", etc. (where "p", "q", and "r" are replaced by clauses, such as "frogs are green" or "Reagan will soon retire"). Linguistically and logically, intentional idioms are a subset of those that Quine calls "referentially opaque". What this means is that many normally valid logical moves are not valid for the clauses "within the scope" of intentional idioms. For instance, normally, if two words happen to be words for the same thing, then one can freely substitute one for the other without affecting the truth of the whole sentence (although one may change its meaning, or effectiveness, or style). Thus, since "Cicero" and "Tully" are names for the same man, from the truth of "Cicero was an orator" we can infer the truth of "Tully was an orator". By contrast, however, the same substitution is not always allowable in "Tom believes that Cicero denounced Catiline", for Tom may believe that Cicero denounced Catiline but not believe that Tully did. So ". . . believes that p" is an opaque idiom.

Clearly this is just another (and more precise) way of putting the point made earlier that intentional relations depend on how their objects are specified. The other points have analogues too. Thus, normally, a relational statement is false if one of the alleged relata doesn't exist; not so within the scope of opaque idioms. And, of course, the fact that the identity of a particular belief depends on the object or objects it is supposed to be about emerges on this treatment as the fact that the ascription of a particular belief depends crucially on the words used in the clause expressing it. We can see now that this condition is essentially the same as the first, and that the second condition (possible non-existence of the object) is also just a special case of opacity: believing that Santa Claus is generous and believing that Poseidon is generous are different beliefs, in spite of the fact that "Santa Claus" and "Poseidon" refer to "the same thing", i.e., to nothing.

Seeing this unity in the various conditions of intentionality is one of the benefits of semantic ascent. Another is that it thus provides us with a relatively formal and uncontroversial test for the intentionality of idioms, and hence a test for appeals to intentionality in a theory. This is an interesting test, for theories relying on intentional idioms-- such as classical "rational agent" economics and cognitive psychology--cannot be formulated in any noncontroversial way within standard logic, while it seems that other theories, pre-eminently in the physical sciences, can be so formulated. The logical oddity of intentional idioms, and their resistance to regimentation, has led Quine and several other theorists to declare the bankruptcy of all intentional theories, on grounds of logical incoherence. The only sound alternatives within the social sciences, then, would have to be theories making no appeal to meaning or intentionality at all: purely behavioristic or purely physiological theories. This claim strikes a familiar note: many psychologists and brain scientists have expressed great skepticism about the utility or permissibility of "mentalist" formulations in their fields--while others of course have held them to be indispensable. The philosophical analysis of intentionality yields a clear logical characterization of this fundamental theoretical division in the social sciences and biology: "mentalist" theories are all and only those making ineliminable use of intentional idioms, and hence inheriting the logical problems of construing those idioms coherently.

Dispensing with intentional theories is not an attractive option, however, for the abstemious behaviorisms and physiological theories so far proposed have signally failed to exhibit the predictive and explanatory power needed to account for the intelligent activities of human beings and other animals. A close examination of the most powerful of these theories reveals intentional idioms inexorably creeping in--for instance in the definition of the stimulus as the "perceived" stimulus and the response as the "intended" effect, or in the reliance on the imputation of "information-bearing" properties to physiological constructs. Moreover the apparent soundness of information-processing theories, and their utility in practice, has strengthened the conviction

that somehow we must be able to make sense of the ineliminable intentional formulations they contain without compromising thoroughgoing materialism.

So we are left with the problem of providing an account of the intentionality of mental states or events which deals with the peculiarities of this feature in a naturalistic way. The most promising approach, found in various versions throughout artificial intelligence and among certain schools of thought in Anglo-American philosophy, is to interpret mental states and processes as representations (perhaps even in a "language of thought"), and to deny the powerful pretheoretical intuition that the intentionality of such mental representations is somehow original or intrinsic--unlike the derivative intentionality of public representations such as sentences, pictures and maps. The familiar and intuitive distinction to be denied is discussed by Haugeland. Our artifacts

Consider an encyclopedia. It has derived intentionality. It contains information about thousands of things in the world, but only insofar as it is a device designed and intended for our use. Suppose we "automate" our encyclopedia, putting all its data into a computer, and turning its index into the basis for an elaborate question-answering system. No longer do we have to look up material in the volumes; we simply type in questions and receive answers. It might seem to naive users as if they were communicating with another person, another entity endowed with original intentionality, but we would know better. A question-answering system is still just a tool, and whatever meaning or aboutness we reside in it is just a byproduct of our practices in using the device to serve our own goals. It has no goals of its own, except for the artificial and derived goal of "understanding" and "answering" our questions correctly.

But suppose we endow our computer with somewhat more autonomous, somewhat less slavish goals. For instance, a chess playing computer has the (artificial, derived) goal of defeating its human opponent, of concealing what it "knows" from us, of tricking us perhaps. But still, surely, it is only our tool or toy, and although many of its internal states have a sort of aboutness or intentionality--e.g., there are states that represent (and hence are about) the current board positions, and processes that investigate (and hence are about) various possible continuations of the game--this is just derived intentionality, not original intentionality.

This persuasive theme (it is not really an argument) has convinced more than a few thinkers that no artifact could have the sort of intentionality we have. Any computer program, any robot we might design and build, no matter how strong the illusion we may create that it has become a genuine agent, could never be a truly autonomous thinker with the same sort of original intentionality we enjoy. The issue can be dramatized with an example.

Consider a standard soft-drink vending machine, designed and built in the United States, and equipped with a transducer device for accepting and rejecting U.S. quarters. Let's call such a device a two-bitser. Normally, when a quarter is inserted into a two-bitser, the two-bitser goes into a state, call it Q, which "means" (note the scare-quotes) "I perceive/accept a genuine U.S. quarter now." Such two-bitserers are quite clever and sophisticated, but hardly foolproof. They do "make mistakes" (more scare-quotes). That is, unmetaphorically, sometimes they go into state Q when a slug or other foreign object is inserted in them, and sometimes they reject perfectly legal quarters--they fail to go into state Q when they are supposed to. No doubt there are detectable patterns in the cases of "misperception". No doubt at least some of the cases of "misidentification" could be predicted by someone with enough knowledge of the relevant laws of physics and design parameters of the two-bitser's transducing machinery, so that it would be just as much a matter of physical law that objects of kind K would put the device into state Q as that quarters would. Objects of kind K would be good "slugs"--reliably "fooling" the transducer.

If objects of kind K became more common in the two-bitser's normal environment, we could expect the owners and designers of two-bitserers to develop more advanced and sensitive

transducers that would reliably discriminate between genuine U.S. quarters and slugs of kind K. Of course trickier counterfeits might then make their appearance, requiring further advances in the detecting transducers, and at some point such escalation of engineering would reach diminishing returns, for there is no such thing as a foolproof mechanism. In the meantime, the engineers and users are wise to make do with standard, rudimentary two-bitser, since it is not cost-effective to protect oneself against negligible abuses.

The only thing that makes the device a quarter-detector rather than a slug-detector or a quarter-or-slug-detector is the shared intention of the device's designers, builders, owners, users. It is only in the environment or context of those users and their intentions that we can single out some of the occasions of state Q as "veridical" and others as "mistaken". It is only relative to that context of intentions that we could justify calling the device a two-bitser in the first place.

This is, apparently, a textbook case of derived intentionality, laid bare. And so of course it embarrasses no one to admit that a particular two-bitser, straight from the American factory and with "Model A Two-Bitser" stamped right on it, might be installed on a Panamanian soft-drink machine, where it proceeded to earn its keep as an acceptor and rejecter of quarter-balboas, legal tender in Panama, and easily distinguished from U.S. quarters by the design and writing stamped on them, but not by their weight, thickness, diameter or material composition.

(I'm not making this up. I have it on excellent authority-- Albert Erler of the Flying Eagle Shoppe, Rare Coins--that Panamanian quarter-balboas minted between 1966 and 1984 are indistinguishable from U.S. quarters by standard vending machines. Small wonder, since they are struck from U. S. quarter stock in American mints. And--to satisfy the curious, although it is strictly irrelevant to the example--the current official exchange rate for the quarter-balboa is indeed \$.25!)

Such a two-bitser, whisked off to Panama (the poor man's Twin-Earth), would still normally go into a certain physical state--the state with the physical features by which we used to identify state Q--whenever a U.S. quarter or an object of kind K or a Panamanian quarter-balboa is inserted in it, but now a different set of such occasions count as the mistakes. In the new environment, U.S. quarters count as slugs, as inducers of error, misperception, misrepresentation, just as much as objects of kind K do. After all, back in the U.S. a Panamanian quarter-balboa is a kind of slug.

Once our two-bitser is resident in Panama, should we say that the state we used to call Q still occurs? The physical state in which the device "accepts" coins still occurs, but should we now say that we should identify it as "realizing" a new state, QB, instead? Well, there is considerable freedom--not to say boredom--about what we should say, since after all a two-bitser is just an artifact, and talking about its perceptions and misperceptions, its veridical and non-veridical states--its intentionality, in short--is "just metaphor". The two-bitser's internal state, call it what you like, doesn't really (originally, intrinsically) mean either "U.S. quarter here now" or "Panamanian quarter-balboa here now". It doesn't really mean anything. So the believers in original intentionality (Fodor, Searle, Dretske, Burge, and Kripke, inter alia) would insist.

The two-bitser was originally designed to be a detector of U.S. quarters. That was its "proper function" (Millikan, 1984), and, quite literally, its *raison d'être*. No one would have bothered bringing it into existence had not this purpose occurred to them. And given that this historical fact about its origin licenses a certain way of speaking, such a device may be primarily or originally characterized as a two-bitser, a thing whose function is to detect quarters, so that relative to that function we can identify both its veridical states and its errors.

This would not prevent a two-bitser from being wrested from its home niche and pressed into service with a new purpose-- whatever new purpose the laws of physics certify it would reliably



serve--as a K-detector, a quarter-balboa detector, a doorstop, a deadly weapon. In its new role there might be a brief period of confusion or indeterminacy. How long a track record must something accumulate before it is no longer a two-bitser, but rather a quarter-balboa-detector (a q-balber)--or a doorstop, or a deadly weapon? On its very debut as a q-balber, after ten years of faithful service as a two-bitser, is its state already a veridical detection of a quarter-balboa, or might there be a sort of force-of-habit error of nostalgia, a mistaken identification of a quarter-balboa as a U.S. quarter?

As described, the two-bitser differs strikingly from us in that it has no provision for memory of its past experiences--or even "memory" (in scare quotes) for its past "experiences". But the latter, at least, could easily be provided, if it was thought to make a difference. To start with the simplest inroad into this topic, suppose the two-bitser (to refer to it by the name of its original baptism) is equipped with a counter, which after ten years of service stands at 1,435,792. Suppose it is not reset to zero during its flight to Panama, so that on its debut there the counter turns over to 1,436,793. Does this tip the balance in favor of the claim that it has not yet switched to the task of correctly identifying quarter-balboas? Would variations and complications on this theme drive your intuitions in different directions?

We can assure ourselves that nothing intrinsic about the two-bitser considered narrowly all by itself and independently of its prior history would distinguish it from a genuine q-balber, made to order on commission from the Panamanian government. Still, given its ancestry, is there not a problem about its function, its purpose, its meaning, on this first occasion when it goes into the state we are tempted to call Q? Is this a case of going into state Q (meaning "U.S. quarter here now") or state QB (meaning "Panamanian quarter-balboa here now")? I would say, along with Millikan (1984), that whether its Panamanian debut counts as going into state Q or state QB depends on whether, in its new niche, it was selected for its capacity to detect quarter-balboas--literally selected, e.g., by the holder of the Panamanian Pepsi-Cola franchise. If it was so selected, then even though its new proprietors might have forgotten to reset its counter, its first "perceptual" act would count as the correct identification of a q-balber, for that is what it would now be for. (It would have acquired quarter-balboa detection as its proper function.) If, on the other hand, the two-bitser was sent to Panama by mistake, or if it arrived by sheer coincidence, its debut would mean nothing, though its utility might soon-- immediately--be recognized and esteemed by the relevant authorities (those who could press it into service in a new role), and thereupon its subsequent states would count as tokens of QB.

Presumably Fodor et al. would be content to let me say this, since, after all, the two-bitser is just an artifact. It has no intrinsic, original intentionality, so there is no "deeper" fact of the matter we might try to uncover. This is just a pragmatic matter of how best to talk, when talking metaphorically and anthropomorphically about the states of the device.

But we part company when I claim to apply precisely the same morals, the same pragmatic rules of interpretation, to the human case. In the case of human beings (at least), Fodor and company are sure that such deeper facts do exist--even if we cannot always find them. That is, they suppose that, independently of the power of any observer or interpreter to discover it, there is always a fact of the matter about what a person (or a person's mental state) really means. Now we might call their shared belief a belief in intrinsic intentionality, or perhaps even objective or real intentionality. There are differences among them about how to characterize, and name, this property of human minds, which I will continue to call original intentionality, but they all agree that minds are unlike the two-bitser in this regard.

I part company with these others, because although they might agree with me (and Millikan) about what one should say in the case of the transported two-bitser, they say that we human beings are not just fancier, more sophisticated two-bitseres. When we say that we go into the state of believing that we are perceiving a U.S. quarter (or some genuine water as opposed to XYZ, or

a genuine twinge of arthritis) this is no metaphor, no mere manner of speaking. A parallel example will sharpen the disagreement:

Suppose some human being, Jones, looks out the window and thereupon goes into the state of thinking he sees a horse (cf. Fodor 1987). There may or may not be a horse out there for him to see, but the fact that he is in the mental state of thinking he sees a horse is not just a matter of interpretation (these others say). Suppose the planet Twin-Earth were just like Earth, save for having schmorses where we have horses. (Schmorses look for all the world like horses, and are well-nigh indistinguishable from horses by all but trained biologists with special apparatus, but they aren't horses, any more than dolphins are fish.) If we whisk Jones off to Twin-Earth, land of the the schmorses, and confront him in the relevant way with a schmorse, then either he really is, still, provoked into the state of believing he sees a horse (a mistaken, non-veridical belief) or he is provoked by that schmorse into believing, for the first time (and veridically), that he is seeing a schmorse. (For the sake of the example, let us suppose that Twin-Earthians call schmorses horses (chevaux, Pferde, etc.) so that what Jones or a native Twin-Earthian says to himself--or others--counts for nothing.) However hard it may be to determine exactly which state he is in, he is really in one or the other (or perhaps he really is in neither, so violently have we assaulted his cognitive system). Anyone who finds this intuition irresistible believes in original intentionality, and has some distinguished company: Fodor, Searle, Dretske, Burge, and Kripke, but also Chisholm 1956, 1957, Nagel 1979, 1986 and Popper and Eccles 1977). Anyone who finds this intuition dubious if not downright dismissible can join me, the Churchlands, Davidson, Haugeland, Millikan, Rorty, Stalnaker, and our distinguished predecessors, Quine and Sellars, in the other corner (along with Douglas Hofstadter, Marvin Minsky and almost everyone else in AI).

There, then, is a fairly major disagreement. Who is right? I have argued in my new book, *The Intentional Stance*, (Dennett, 1987) that clinging to the doctrine of original intentionality is the primary source of perplexity in contemporary Anglo-American philosophy of mind, but the arguments for that broad indictment cannot be presented in the short compass of this introductory talk. I can, however, give a summary expression of the main positive point of my theory: the intentionality of our mental states and processes is derived in just the same way as that of our books and maps (and the inner states of our robots). Suppose you have composed a shopping list, on a piece of paper, to guide your shopping behavior. The marks on the piece of paper have derived intentionality, of course, but if you forgo the shopping list and just remember the wanted items in your head, whatever it is that "stores" or "represents" the items to be purchased in your brain has exactly the same status as the trails of ink on the paper. There is no more real, or intrinsic, or original intentionality than that.