

Хилари Патнэм

**Разум, истина
и история**

• • • • •



Данное издание выпущено в рамках программы Центрально-Европейского Университета «Translation Project» при поддержке Регионального издательского центра Института «Открытое общество» (OSI –Budapest) и Института «Открытое общество» (Фонд Сороса) – Россия

Патнем Хилари

П 202 Разум, истина и история / Пер. с англ. Т. А. Дмитриева, М. В. Лебедева. — М.: Праксис, 2002. — 296 с. — (Серия «Философия»).

ISBN 5-901574-09-5

Работа известного американского философа, выдержавшая после своей публикации в свет в 1981 году множество переизданий, посвящена исследованию основных проблем современной философии. Автор подробно рассматривает различные проблемы, связанные с теорией референции, философией сознания, познанием внешнего мира, с основными формами использования языка в философии, а также проблему соотношения фактов и ценностей в гуманитарном познании.

Книга будет интересна философам, социологам, политологам, а также всем, кто интересуется современной философией.

ББК 87.3

© Hilary Putnam 1981. Reason, Truth and History. Was originally published in 1981. This translation is published by arrangement with Cambridge University Press.

© Т. А. Дмитриев, М. В. Лебедев,
пер. с англ., 2002

© М. В. Лебедев, послесловие, 2002

© А. Кулагин, А. Эльконин.
оформление обложки, 2002

© Издательская группа «Праксис», 2002

ISBN 5-901574-09-5

СОДЕРЖАНИЕ

Предисловие	9
1. Мозги в бочке	14
2. Проблема референции	38
3. Две философские перспективы	70
4. Сознание и тело	103
5. Две концепции рациональности	139
6. Факт и ценность	171
7. Разум и история	198
8. Вклад науки в современные концепции рациональности	231
9. Ценности, факты и познание	262
Приложение	281
<i>М. В. Лебедев. «The Matrix never has you». Хилари Патнэм и традиционная философия</i>	283

УВЕДОМЛЕНИЕ

Читатель, не слишком сведущий в технических тонкостях философии, или же желающий составить общее представление о проблемах, которые обсуждаются в этой книге, вполне может начать чтение с главы 5 и, дочитав книгу до конца, вернуться затем к знакомству с главами 1—4.

ПРЕДИСЛОВИЕ

Цель, которую я ставил перед собой в данной работе, заключается в том, чтобы ослабить то удушающее влияние, которое ряд дихотомий, как мне кажется, оказывает на мышление и философов, и обывателей. Главная из них — дихотомия между объективными и субъективными взглядами на истину и на разум. Явление, о котором я думаю, состоит в следующем: как только такая дихотомия, как дихотомия между «объективным» и «субъективным» принята, и принята не как простая пара категорий, но как характеристика типов подходов и стилей мышления, мыслители начинают рассматривать термины этой дихотомии почти как идеологические ярлыки. Сегодня многие философы, а возможно, большинство, поддерживают некоторую версию «копирующей» теории истины — т. е. концепцию, согласно которой утверждение является истинным в том случае, если оно «соответствует (независимым от сознания) фактам». Философы этой фракции видят единственную альтернативу в опровержении объективности истины и в капитуляции перед идеей о том, что все схемы мышления и все точки зрения безнадежно субъективны. Дерзкое меньшинство неизбежным образом (Кун, по крайней мере иногда; Фейерабенд и такой выдающийся континентальный философ, как Фуко) занимает противоположную позицию. Они *соглашаются*, что альтернатива наивной копирующей концепции истины должна рассматривать системы мышления, идеологии и даже (в случае Куна и Фейерабенда) научные теории как субъективные, и они *продолжают энергично выдвигать* релятивистский и субъективный подход.

То, что философский спор в известном смысле принимает характер идеологического спора, — само по себе не обязательно *плохо*: новые идеи, даже в наиболее точных науках, часто и поддерживаются, и атакуются с пристрастной энергией. Даже в политике поляризация и идеологический пыл иногда необходимы, чтобы привнести в проблему моральную серьезность. Но со временем и в философии, и в политике новые идеи становятся старыми идеями; то, что некогда было стимулирующим,

становится предсказуемым и раздражающим; и то, что некогда служило сосредоточению внимания там, где оно должно быть сосредоточено, позже препятствует рассмотрению новых альтернатив. Это и случилось теперь в дебатах между сторонниками корреспондентных подходов к проблеме истины и субъективистских подходов. В первых трех главах этой книги я попытаюсь объяснить концепцию истины, которая объединяет объективные и субъективные компоненты. Этот подход возвращается, по крайней мере по духу, к идеям Иммануила Канта. Согласно этому подходу, мы можем отклонять наивную «копирующую» концепцию истины без того, чтобы иметь необходимость считать, что истина — это вопрос *Zeitgeist*¹ или вопрос «гештальных переключателей», или вопрос идеологии.

Подход, который я буду защищать, состоит, грубо говоря, в том, что существует чрезвычайно близкая связь между понятиями *истины* и *рациональности*; что, выражаясь еще более приблизительно, единственный критерий для того, что является фактом — то, что это *рационально* приемлемо. (Я говорю об этом в совершенно буквальном и общем смысле; таким образом, если может быть рационально принято, что картина красива, то это может быть *факт*, что картина является красивой.) Согласно этой концепции, могут быть *ценностные факты*. Но отношение между рациональной приемлемостью и истиной — это отношение между двумя различными понятиями. Утверждение может быть рационально приемлемо *в некоторое время*, но не *истинно*, и эта реалистичная интуиция будет сохранена в моей теории.

Я не верю, однако, что рациональность определяется множеством неизменяющихся «канонов» или «принципов»; методологические принципы связаны с нашим видением мира, включая наше видение нас самих как части мира, и подвержены изменениям со временем. Таким образом, я *соглашаюсь* с философами-субъективистами в том, что не существует никакого установленного, внеисторического *органа*, который определяет, что именно является рациональным; но из того факта, что наши концепции разума развиваются с ходом истории, я не заключаю, что сам разум вообще может быть *чем-то* (или развиваться в нечто). И при этом я не прихожу к некоторой причудливой смеси

1. Дух времени (нем.). — Прим. пер.

культурного релятивизма и «структурализма», подобно французским философам. Дихотомия: либо внеисторические, неизменяющиеся каноны рациональности, *либо* культурный релятивизм — это дихотомия, которую я расцениваю как устаревшую.

Другая особенность этого подхода состоит в том, что рациональность не ограничена лабораторной наукой и не различается фундаментальным образом в лабораторной науке и вне нее. Подобная концепция представляется мне пережитком позитивизма; пережитком идеи о том, что научный мир некоторым способом построен из «чувственных данных» и идеи о том, что термины лабораторных наук «определяются операционально». Я не стану посвящать много места критике операционалистских и позитивистских подходов в науке; они достаточно подробно критиковались в последние примерно двадцать лет. Но эмпиристская идея о том, что «чувственные данные» составляют некоторый объективный «фундамент» для по крайней мере части нашего знания, будет повторно исследована в свете того, что мы должны сказать об истине и рациональности (в главе 3).

Короче говоря, я предложу подход, согласно которому сознание не просто «копирует» мир, подлежащий описанию в соответствии с Единой Истинной Теорией. Однако мой подход — это и не такой подход, при котором сознание *составляет* мир (или составляет его в рамках ограничений, налагаемых «методологическими канонами» и независимыми от сознания «чувственными данными»). Если использовать метафорический язык, то метафора будет такой: сознание и мир совместно составляют сознание и мир. (Или, делая эту метафору еще более гегельянской, Универсум составляет Универсум — вместе с сознаниями — коллективно — играя особую роль в этом составлении.)

Заключительная особенность моей теории рациональности состоит вот в чем: я попытаюсь показать, что наше понятие рациональности, в своем основании, представляет собой лишь часть нашей концепции процветания человека, нашей идеи блага. Истина глубоко зависит от того, что недавно было названо «ценностями» (глава 6). И то, что мы сказали выше о рациональности и истории, также применимо к ценностям и истории; нет никакого данного, внеисторического множества «моральных принципов», которые определяют раз и навсегда, в чем состоит процветание человека; но это не подразумевает, что все это

просто относительно и принадлежит культуре. Поскольку нынешнее положение дел в теории истины — нынешняя дихотомия между копирующими теориями истины и субъективными теориями истины — является по крайней мере частично ответственной, согласно моему подходу, за печально известную дихотомию «факт/ценность», мы можем выйти за пределы последней, лишь перейдя на очень глубокий уровень и исправляя наши теории истины и рациональности. (Дихотомия «факт/ценность», как она традиционно понимается, фактически обязывает своих приверженцев принять вместе с ней некоторый вид релятивизма). Нынешние подходы к проблеме истины отчуждены; они заставляют своих приверженцев терять ту или иную часть своего «я» и мира, видеть мир как просто состоящий из элементарных частиц, отклоняющихся в пустоте («физикалистский» подход, который видит *научное* описание как сведение к Единой Истинной Теории), или видеть мир как просто состоящий из «действительных и возможных чувственных данных» (старый эмпиристский подход), или вообще отрицать, что есть некоторый мир, противопоставленный множеству историй, которые мы составляем по различным (главным образом бессознательным) причинам. И моя цель в этой работе состоит в том, чтобы дать набросок ведущих идей неотчужденного подхода.

Моя Спенсеровская лекция, «Философы и человеческое понимание» (прочитанная в Оксфордском университете в 1979 году) накладывается на существующий текст, став промежуточным результатом текущей работы, равно как и статья ««Si Dieu est mort, alors tout est permis»... (réflexions sur la philosophie du langage)», *Critique*, 1980.

Исследовательский грант от Национального Научного фонда поддержал исследование, связанное с этой книгой, в течение 1978—80 годов². Я с благодарностью признаю эту поддержку.

Томас Кун и Рут Анна Патнэм изучили черновики этой книги и дали мне компетентную критику и мудрые советы. Мне помогли также советом и критикой многие друзья, включая Неда Блока, Дэвида Хелмана и Джастина Лейбера, а также студенты на

2. «Оценка научных теорий: исторический подход против формальных методологических подходов»; проект № SOC78-04276.

моих различных лекциях и семинарах в Гарварде. Несколько глав читались как лекции в Лиме весной 1980 года (поездка, сделанная возможной грантом от Комиссии Фулбрайт), а глава 2 была действительно закончена в течение моего пребывания в Лиме. Я извлек пользу в этот период от обсуждений с Леопольдо Кьяппо, Алберто Кордеро Леккой, Энрике Фернандесом, Франсиско Миро Квесадой и Хорхе Секадой. Книга целиком (в более ранней версии) читалась как лекции в университете Франкфурта летом 1980 года, и я благодарен моим тамошним коллегам (особенно Вильгельму Эсслеру и Райнеру Траппу), моей очень стимулирующей группе студентов и другим моим друзьям в Германии (особенно Дитеру Хенриху, Манон Фассбиндер и Вольфгангу Штегмюллеру) за ободрение и плодотворные обсуждения.

Все мои коллеги по философскому факультету Гарварда заслуживают быть отмеченными для индивидуальной благодарности. В последние годы Нельсон Гудмен и я обнаружили сходство в наших подходах, и в то время как первый черновик этой книги был написан прежде, чем я имел возможность ознакомиться с его *«Способами создания миров»*, читать его текст и обсуждать с ним эти проблемы было очень важно для меня на многих этапах.

Я также благодарен Джереми Майнотту за поддержку и советы в качестве редактора.

1. МОЗГИ В БОЧКЕ

Муравей ползет по песку и оставляет на нем след. Совершенно случайно этот след таким образом изгибается и перекрещивается, что образует узнаваемую карикатуру на Уинстона Черчилля. Нарисовал ли муравей картинку Уинстона Черчилля — картинку, которая *изображает* Черчилля?

Большинство людей, не слишком вдаваясь в размышления, ответит отрицательно. В конце концов, муравей никогда не видел Черчилля или хотя бы изображения Черчилля; у него не было намерения изображать Черчилля. Он просто провел линию (и даже *это* непреднамеренно), которую мы можем «рассматривать как» изображение Черчилля.

Мы можем выразить это, сказав, что линия «сама по себе» не репрезентирует¹ что бы то ни было в большей степени, чем нечто иное. Сходства (некоторого очень сложного вида) с чертами Уинстона Черчилля еще недостаточно для того, чтобы позволить чему-то репрезентировать Уинстона Черчилля или осуществлять к нему референцию. Но сходство и не является для этого необходимым: в нашем сообществе напечатанные значки «Уинстон Черчилль», произнесенные слова «Уинстон Черчилль» и многие другие вещи используются для репрезентации Черчилля (хотя и не изобразительной), не обладая при этом сходством с Черчиллем того же вида, что и изображение, даже абрисный рисунок. Если *сходство* не является необходимым или

1. В этой книге термины «репрезентация» и «референция» всегда указывают на отношение между словом (или другими видами знаков, символов или представлений) и чем-то, что существует в действительности (т. е. не является лишь «мысленным объектом»). Есть некоторый смысл, в котором я могу «указывать» (осуществлять референцию) на то, что не существует, но это не тот смысл, в котором я здесь употребляю слово «референция». Раньше то, что я называю «репрезентацией» и «референцией», называли *денотацией*.

Кроме того, я следую образцу современных логиков и использую слово «существует», подразумевая «существует в прошлом, настоящем или будущем». Таким образом, Уинстон Черчилль «существует», и мы можем «репрезентировать» Уинстона Черчилля или «осуществлять референцию» к нему даже несмотря на то, что его нет в живых.

достаточным для того, чтобы сделать нечто репрезентацией чего-то еще, то как вообще что бы то ни было может быть необходимым или достаточным для этой цели? Как вообще одна вещь может репрезентировать («представлять» и т. п.) другую, отличную от нее?

Ответ может показаться легким. Предположим, что муравей видел Уинстона Черчилля, и предположим, что у него достаточно ума и способностей, чтобы его нарисовать. Предположим, что он создал эту карикатуру *намеренно*. Тогда эта линия действительно представляла бы Черчилля.

С другой стороны, предположим, что линия имела бы форму УИНСТОН ЧЕРЧИЛЬ. И предположим, что это произошло бы совершенно случайно (не станем обращать здесь внимание на невероятность такого события). Тогда «печатные значки» УИНСТОН ЧЕРЧИЛЬ *не* репрезентировали бы Черчилля, хотя эти же печатные значки репрезентируют Черчилля практически в каждой книге, где они сегодня встречаются.

Поэтому может показаться, что для репрезентации необходимо (или главным образом необходимо) *намерение*.

Но для того, чтобы иметь намерение, что *что бы то ни было*, даже индивидуальный язык (и даже слова «Уинстон Черчилль», произнесенные мною про себя, а не вслух), должно *репрезентировать* Черчилля, я должен прежде всего быть способным *подумать* о Черчилле. Если линии на песке, звуки и т. п. не могут ничего репрезентировать «сами по себе», то как могут что-то репрезентировать «сами по себе» мысленные формы? Или все-таки могут? Как мышление может обратиться к чему-то внешнему и ухватить его?

В прошлом некоторые философы перескакивали от такого рассмотрения к тому, что они считали доказательством *сущностно не-физической природы сознания*. Аргумент здесь прост: сказанное нами о муравьиной кривой применимо к любому физическому объекту. Ни один физический объект не может сам по себе указывать на одну вещь скорее, чем на другую; тем не менее, *мысли в сознании* очевидно преуспевают в указании на одну вещь скорее, чем на другую. Поэтому мысли (и, соответственно, сознание) по природе своей сущностно отличаются от физических объектов. Мысли обладают характеристикой *интенциональности* — они могут указывать на нечто иное; ничто физическое

не обладает «интенциональностью», не считая того обстоятельства, что интенциональность производна от некоторого употребления этой физической вещи сознанием (или, по крайней мере, так признается). Однако это слишком торопливое решение; простое постулирование таинственных сил сознания само по себе ничего не дает. Но проблема вполне реальна. Как возможна интенциональность, референция?

МАГИЧЕСКИЕ ТЕОРИИ РЕФЕРЕНЦИИ

Мы видели, что созданная муравьем «картинка» не имеет никакой необходимой связи с Уинстоном Черчиллем. Тот факт, что «картинка» имеет определенное «сходство» с Черчиллем, сам по себе не делает ее настоящей картинкой или не делает ее репрезентацией Черчилля. Если только муравей не является разумным муравьем (каковым он не является) и не знает Черчилля (а он не знает), начерченная им кривая — не картинка и даже не репрезентация чего бы то ни было. Некоторые примитивные народы верят, что репрезентации (в особенности *имена*) связаны со своими носителями с необходимостью; что знание «подлинного имени» кого-либо или чего-либо дает подлинную власть над этим человеком или вещью. Эта власть происходит из *волшебной связи* между именем и его носителем; когда же мы понимаем, что имена связаны со своими носителями *лишь* контекстуально, случайно, конвенционально, нам уже трудно увидеть, каким образом знание имени может иметь какое бы то ни было мистическое значение.

Важно понять, что для физических изображений действуют те же правила, что и для ментальных образов и вообще для ментальных репрезентаций; ментальные репрезентации связаны с тем, что они представляют, не с большей необходимостью, чем физические. Обратное предположение — пережиток магического мышления.

Возможно, легче понять это замечание в случае с ментальными *образами*. (Возможно, первым философом, ухватившим огромное значение этого замечания, даже если он не первым его обнаружил, был Витгенштейн). Предположим, что где-то есть планета, населенная человеческими существами — появившимися там в результате эволюции, или высадки инопланетян, или

как бы то ни было. Предположим, что эти люди — во всем остальном полностью похожие на нас — никогда не видели *деревьев*. Предположим, что они никогда не представляли деревьев (возможно, растительность существует на их планете лишь в виде плесени). Предположим, что однажды на их планету случайно попала картинка, изображающая дерево; ее обронил пролетающий мимо космический корабль, никак не вступивший с ними в контакт. Представьте себе, как они ломают голову над этой картинкой. Что бы это могло быть? Им приходят на ум всевозможные предположения: может быть, это какое-то строение? Балдахин? Или какое-то животное? Но они, как мы продолжаем предполагать, не приближаются к истине.

Для нас картинка является репрезентацией дерева. Для этих же людей картинка представляет лишь некоторый странный объект, чьи природа и назначение неизвестны. Предположим, что в результате взгляда на картинку у одного из них возник ментальный образ, в точности подобный одному из моих ментальных образов дерева. Однако его ментальный образ не является *репрезентацией дерева*. Это лишь репрезентация странного предмета (чем бы он ни был), который представляет таинственная картинка.

Здесь можно возразить, что этот ментальный образ *в действительности* является репрезентацией дерева при условии, что картинка, вызвавшая этот ментальный образ, сама изначально была репрезентацией дерева. Существует каузальная цепь, ведущая от действительных деревьев к ментальному образу, даже если эта цепь весьма необычна.

Но можно представить себе ситуацию, где даже и эта каузальная цепь отсутствует. Предположим, что «картинка дерева», которую обронил космический корабль, была на самом деле не картинкой, изображающей дерево, а случайным результатом разлития красок. Даже если бы этот результат выглядел в точности как картина дерева, он поистине не в большей степени изображал бы дерево, чем муравьиная «карикатура» на Черчилля изображала бы Черчилля. Мы даже можем представить себе, что космический корабль, обронивший картинку, прилетел с планеты, где тоже ничего не знают о деревьях. Тогда у рассматривающих картинку людей возникли бы ментальные образы, количественно идентичные моему образу дерева, но эти обра-

зы не представляли бы дерево в большей степени, нежели что-то другое.

То же самое относится и к *словам*. Ряд значков на бумаге может быть прекрасным описанием деревьев, но если он был произведен обезьянами, наугад нажимающими клавиши на машинке в течение миллионов лет, то эти слова ни на что не указывают. И если бы кто-то запомнил эти слова и произнес их про себя, не понимая, то они ни на что не указывали бы, будучи помысленными.

Представим себе, что человек, произносящий эти слова про себя, оказался загипнотизирован. Предположим, что эти слова — на японском, а нашему герою внушили, что он понимает японский. Предположим, что по мере того, как он мыслит эти слова, у него появляется «чувство понимания». (Хотя если кто-нибудь вмешается в ход его мысли и спросит его, что *значат* те слова, которые он мыслит, то он обнаружит, что не может ответить.) Возможно, иллюзия нашего героя будет настолько полной, что он сможет обмануть японца-телепата. Но если он не мог бы использовать слова в правильных контекстах, отвечать на вопросы о том, что он «думает» и т. д., то он не понимал бы их.

Комбинируя рассказанные мною научно-фантастические истории, мы можем выдумать такой случай, когда некто произносит в уме слова, которые фактически являются описаниями слов в некотором языке, и одновременно у него возникают соответствующие ментальные образы, но при этом он *не* понимает эти слова и *не* знает, что такое дерево. Мы даже можем представить себе, что ментальные образы были вызваны разлитием красок (хотя наш герой под воздействием гипноза думает, что это — образы чего-то соответствующего его мысли, но не может ответить на вопрос о том, что же это такое). И мы можем представить, что язык, на котором думает наш герой — это такой язык, о котором ни гипнотизер, ни гипнотизируемый никогда и не слышали. Возможно, эти «бессмысленные предложения», какими их полагает гипнотизер, являются описаниями деревьев на японском лишь в силу простого совпадения. Короче говоря, все происходящее в сознании нашего героя может быть количественно идентично тому, что происходит в сознании носителя японского языка, который *действительно* думает о деревьях — но ничто из этого не будет указывать на деревья.

Конечно же, все это в действительности невозможно, в том смысле, что в действительности невозможно, чтобы обезьяна случайно напечатала «Гамлета». Иными словами, обратная вероятность столь велика, что это никогда не произойдет в действительности (как мы полагаем). Но это не является логически невозможным или даже физически невозможным. Это *могло бы* произойти (в соответствии с физическими законами и, возможно, с действительным состоянием вселенной, если в ней много планет с разумными существами). И если бы это произошло, это было бы поразительное проявление важной концептуальной истины: что даже обширная и сложная система репрезентаций, как словесных, так и визуальных, не имеет *внутренней*, встроенной, волшебной связи с тем, что она представляет — связи, независимой от того, как она причинно обусловлена, и от того, что каковы диспозиции говорящего или думающего. И это справедливо и в том случае, когда система репрезентаций (слов и образов, в случае наших примеров) воплощена физически — слова написаны или произнесены, а картинки являются физическими изображениями, — и в том, когда она воплощена лишь в сознании. Помысленные слова и ментальные картины не представляют *внутренне* то, на что они указывают.

ДЕЛО О МОЗГАХ В БОЧКЕ

Одна из научно-фантастических возможностей, обсуждаемых философами, такова. Представим себе, что человеческое существо (допустим, вы сами) подверглось операции, сделанной злодеем-ученым. Ваш мозг был отделен от тела и помещен в бочку с питательным раствором, поддерживающим этот мозг живым. Нервные окончания были присоединены к некоему супернаучному компьютеру, поддерживающему у той личности, которой принадлежит этот мозг, полную иллюзию, что все в порядке. Как будто вокруг люди, предметы, небо и т. д.; но в действительности все, что ощущает наш герой, то есть вы, является результатом электронных импульсов, передающихся от компьютера к нервным окончаниям. Компьютер такой умный, что если наш герой пытается поднять руку, то реакция компьютера послужит причиной того, что он «увидит» и «почувствует», как рука поднимается. Более того, меняя программу, злодей-ученый может

заставить жертву «почувствовать» (или прогаллюцинировать) любую ситуацию или окружение, какие хочет. Он может также стереть память об операции на мозге, в результате чего жертва будет считать, что всегда была в таком окружении. Жертве даже может казаться, что она сейчас сидит и читает вот эти слова о забавном, но совершенно абсурдном предположении, что есть такой злодей-ученый, который удаляет людям мозги и помещает их в бочки с питательным раствором, поддерживающим эти мозги живыми. При этом нервные окончания присоединены к некоему супернаучному компьютеру, поддерживающему у той личности, которой принадлежит этот мозг, полную иллюзию, что...

Когда такая возможность упоминается на лекциях по теории познания, то цель, конечно, состоит в том, чтобы современными средствами поставить классическую проблему скептицизма по отношению к внешнему миру. (*Откуда вы знаете, что вы сами не находитесь в этом затруднении?*) Но это затруднение также полезно в качестве инструмента для постановки вопросов, касающихся отношения между сознанием и миром.

Вместо одного-единственного мозга в бочке теперь представим себе, что все человеческие существа (возможно, все существа, наделенные ощущениями) представляют собой мозги в бочке (или нервные системы в бочке для тех существ с минимальными нервными системами, которые уже считаются «чувствующими»). Конечно же, злодей-ученый должен находиться снаружи. Или не должен? А может быть, нет никакого злодея-ученого, а просто вселенная (хотя это и абсурдно) состоит из автоматических механизмов, обслуживающих бочку с мозгами и нервными системами?

На этот раз предположим, что автоматические механизмы запрограммированы так, чтобы дать нам *массовую* галлюцинацию, а не множество отдельных, не связанных галлюцинаций. Таким образом, когда мне кажется, что я разговариваю с вами, вам кажется, что вы слышите мои слова. Конечно, это не значит, что мои слова действительно достигают ваших ушей — поскольку у вас (в действительности) нет ушей, а у меня рта и языка. Скорее, когда я порожаю слова, происходит следующее: центробежные импульсы проходят от моего мозга к компьютеру, что заставляет меня «слышать» мой собственный голос, про-

износящий эти слова, «чувствовать», как мой язык двигается, и т. д., а вас заставляет «слышать» мои слова, «видеть», как я говорю и т. д. В этом случае мы, в некотором смысле, действительно общаемся. Я не заблуждаюсь относительно вашего действительного существования (только относительно существования вашего тела и «внешнего мира», отдельно от мозгов). С определенной точки зрения, даже неважно, что «целый мир» — это лишь коллективная галлюцинация; поскольку вы, в конце концов, действительно слышите мои слова, когда я говорю с вами, даже если механизм этого процесса не таков, как мы предполагаем. (Конечно, если бы мы были влюбленными, занимающимися любовью, а не просто двумя собеседниками, то предположение о том, что мы всего лишь два мозга в бочке, разочаровало бы нас).

Теперь я хочу задать вопрос, который покажется очень глупым и очевидным (по крайней мере некоторым людям, включая некоторых весьма изощренных философов), но который довольно быстро заведет нас в настоящие философские бездны. Предположим, что вся эта история — действительно правда. Могли бы мы, если бы мы были такими мозгами в бочке, *сказать* или *подумать*, что мы таковы?

Я буду доказывать, что ответ на этот вопрос отрицательный. Фактически я буду утверждать, что предположение о том, что мы на самом деле мозги в бочке, не может быть истинным, хотя оно не нарушает физических законов и вполне совместимо со всем нашим опытом. *Оно не может быть истинно*, потому что в определенном смысле оно само себя отрицает.

Аргумент, который я здесь хочу представить, необычен, и мне понадобилось несколько лет, чтобы убедить себя в его правильности. Но это корректный аргумент. Станным его заставляет выглядеть то обстоятельство, что он оказывается связанным с некоторыми самыми глубокими философскими вопросами. (Он впервые пришел мне в голову, когда я размышлял над одной из теорем современной логики — теоремой Сколема-Левенгейма, и внезапно увидел связь между этой теоремой и некоторыми аргументами «Философских исследований» Витгенштейна.)

«Самоотрицающее предположение» — это такое утверждение, истинность которого предполагает его же ложность. Например, рассмотрим тезис «Все общие утверждения ложны». Это общее

утверждение. Тогда, если оно истинно, то оно должно быть ложно. Следовательно, оно ложно. Иногда тезис называют самоотрицающим, если его ложность вызывается *предположением, что тезис вообще поддержан или провозглашен*. Например, тезис «Я не существую» является самоотрицающим (для любого «я»). Поэтому если кто-то думает о том, что он существует, то можно быть уверенным, что он существует (как утверждал Декарт).

Я покажу, что предположение о том, что мы представляем собой мозги в бочке, обладает именно этим свойством. Если мы вообще можем оценивать его с точки зрения истинности или ложности, то оно не истинно (как я покажу). Следовательно, оно не истинно.

Прежде, чем я приведу аргумент, рассмотрим, почему нам кажется таким странным то обстоятельство, что такой аргумент вообще может быть дан (по крайней мере некоторым философам — приверженцам концепции истины как подобия). Мы признали, что то положение, что существует мир, где все наделенные ощущениями существа представляют собой мозги в бочке, совместимо с законами физики. Как выражаются философы, есть «возможный мир», где все наделенные ощущениями существа представляют собой мозги в бочке. (Рассуждение о «возможном мире» таково, что оно как бы предполагает существование *места*, где всякое абсурдное предположение истинно; поэтому в философии оно может вводить в заблуждение.) Человеческие существа в этом возможном мире обладают теми же знаниями, что и *мы*. У них те же мысли (по крайней мере, через их сознание проходят те же слова, образы, формы мысли и т. д.). Однако я утверждаю, что мы можем дать аргумент, показывающий, что мы — не мозги в бочке. Как возможен такой аргумент? И почему его не могли бы точно так же выдвинуть люди в возможном мире, действительно *являющиеся* мозгами в бочке?

Ответ будет (вкратце) таков: хотя люди в этом возможном мире могут мыслить и «произносить» все слова, которые мы мыслим и произносим, они не могут (как я утверждаю) *указывать* на то же, на что и мы. В частности, они не могут помыслить или сказать, что они — мозги в бочке (*даже думая «мы — мозги в бочке»*).

ТЕСТ ТЬЮРИНГА

Предположим, что кому-то удалось изобрести компьютер, который может поддерживать осмысленную беседу с человеком (на столько тем, на сколько мог бы разумный человек). Как можно решить, «разумен» ли компьютер?

Английский логик Алан Тьюринг предложил следующий тест²: пусть некто поддерживает разговор с компьютером и с неизвестным человеком. Если испытатель не сможет различить, где компьютер, а где человек (предположим, что тест повторяется достаточное количество раз с разными собеседниками), то компьютер разумен. Вкратце, вычислительная машина разумна, если она может пройти «тест Тьюринга». (Разговор ведется, разумеется, не лично, поскольку испытатель не должен знать облик своих собеседников. Нельзя использовать и голос, так как механический голос может просто звучать иначе, чем человеческий. Скорее представим себе, что беседа ведется с помощью электрической пишущей машинки. Испытатель печатает на ней утверждения, вопросы и т. п., а его партнеры — машина и человек — отвечают с помощью электрической клавиатуры. Кроме того, машина может *лгать* — на вопрос «Ты машина?» она может ответить: «Нет, я тут лаборант».)

Предложение такого теста в качестве решающего теста на разумность было подвергнуто критике многими авторами (в принципе никоим образом не враждебными идее о том, что машина может быть разумной). Но не это сейчас занимает нас. Я хочу использовать общую идею теста Тьюринга, общую идею *диалогического теста на компетентность*, для другой цели — цели исследования понятия *референции*.

Представим ситуацию, в которой задача состоит не в том, чтобы определить, является ли собеседник человеком или машиной, а в том, чтобы определить, использует ли он слова с той же референцией, что и мы. Очевидно, тест состоял бы в том, чтобы так же вести разговор и, если при этом не возникает проблем, если собеседник оказывается неотличимым от кого-то, о ком нам заранее достоверно известно, что он говорит на том

2. Turing A. M. Computing Machinery and Intelligence // Mind, 1950. Перепечатано в: A. R. Anderson (ed.). Minds and Machines.

же языке, то мы заключаем, что собеседник указывает на предметы тем же способом, что и мы. В тех случаях, когда тест Тьюринга направлен описанным образом на определение существования (общей, разделяемой) референции, я буду называть его *Тестом Тьюринга на Референцию*. Подобно тому, как философы обсуждали вопрос, является ли изначальный тест Тьюринга *решающим* тестом на разумность — т. е. вопрос о том, является ли машина, проходящая тест не один раз, а регулярно, разумной *с необходимостью*, — я хочу обсудить вопрос о том, является ли предложенный Тест Тьюринга на Референцию *решающим* тестом на общую референцию.

Ответ окажется отрицательным. Тест Тьюринга на Референцию не является решающим. На практике это, конечно, превосходный способ проверки; однако логически возможно (хотя, конечно, крайне маловероятно), чтобы кто-то прошел Тест Тьюринга на Референцию, ни на что не указывая. Отсюда следует, как мы увидим, что мы можем расширить наше наблюдение о том, что слова (и целые тексты и дискурсы) не связаны со своими референтами *с необходимостью*. Даже если мы рассматриваем не слова сами по себе, а правила, определяющие, какие слова могут быть надлежащим образом произведены в данном контексте — даже если мы рассматриваем, выражаясь компьютерным жаргоном, *программы для использования слов*, — здесь все еще не появится определенная референция, если только эти программы сами не *указывают на нечто внеязыковое*. Это будет решающим шагом в процессе достижения заключения о том, что Обитатели Мира Мозгов-в-Бочке вообще не могут указывать на что-либо внешнее (и тем самым не могут сказать, что они — Обитатели Мира Мозгов-в-Бочке).

Предположим, например, что я нахожусь в Тьюринг-ситуации (веду «имитационную игру», в терминологии Тьюринга) и мой собеседник в действительности машина. Предположим, что машина запрограммирована так, что способна выиграть («пройти» тест). Представим, что машина запрограммирована давать прекрасные ответы по-английски на утверждения, вопросы, замечания и т. д. на английском языке, но у машины нет ни органов чувств (кроме подсоединения к электрической пишущей машинке), ни органов движения (кроме электрической пишущей машинки). (Насколько я себе представляю, Тьюринг не ут-

верждает, что обладание органами чувств или движения необходимо для наличия сознания или разума). Допустим, что не только у машины нет электронных ушей, глаз и т. д., но также в ее программе для ведения имитационной игры отсутствуют соответствующие модули для введения данных с таких органов чувств или для контроля над телом. Что мы скажем о такой машине?

Для меня очевидно, что мы не можем и не должны приписывать референции такому устройству. В самом деле, машина может прекрасно рассуждать, скажем, о видах Новой Англии, но она не может распознать яблоню или яблоко, гору или корову, поле или колокольню, если бы они находились перед ней.

Мы получили машину для производства предложений в ответ на предложения. Но ни одно из этих предложений никак не связано с реальным миром. *Если соединить две машины и заставить их играть в имитационную игру друг с другом, то они будут «обманывать» друг друга бесконечно, даже если весь остальной мир исчезнет!* К тому, чтобы рассматривать речь машины о яблоках как указывающую на яблоки реального мира, не больше оснований, чем рассматривать «рисунок» муравья» как указывающий на Уинстона Черчилля.

Иллюзию референции, значения, разумности и т. д. здесь производит тот факт, что речь машины указывает на яблоки, колокольни, Новую Англию и т. д. в пределах нашей конвенции по поводу репрезентации. Точно так же может существовать иллюзия, что муравей изобразил Черчилля — по той же самой причине. Но мы можем воспринимать яблоки и поля, воздействовать на них и взаимодействовать с ними. Наша речь о яблоках и полях тесно связана с невербальными отношениями между нами и этими предметами. Есть «правила языкового входа», которые ведут нас от знаний о яблоках к таким произнесениям, как «Я вижу яблоко», и «правила языкового выхода», которые ведут нас от выраженных в языковой форме решений («Куплю яблок») к действиям — другим действиям, нежели говорение. В отсутствие либо правил языкового входа, либо правил языкового выхода нет причин рассматривать речь машины (или двух машин, если мы берем случай с двумя машинами, играющими в имитационную игру друг с другом) как нечто большее, чем синтаксическая игра. Синтаксическая игра, которая *походит* на осмысленную

речь, — пожалуй, но лишь в той степени (и не в большей), в какой муравьиная кривая походит на язвительную карикатуру.

В случае с муравьем мы могли бы привести такой аргумент: муравей провел бы ту же кривую, если бы Черчилль никогда не существовал. В случае с машиной мы не можем рассуждать подобным образом; если бы яблоки, деревья, колокольни и поля не существовали, то, по всей вероятности, программисты не создали бы именно эту программу. Хотя машина не *воспринимает* яблоки, поля или колокольни, ее создатели-проектировщики воспринимали все это. Существует *некоторая* каузальная связь между машиной и яблоками и т. д. реального мира — связь через перцептуальный опыт и знания создателей-проектировщиков. Но такая слабая связь едва ли достаточна для референции. Прежде всего, логически возможно, хотя и фантастически невероятно, что та же самая машина *могла бы* существовать и в том случае, если бы яблоки, поля и колокольни не существовали. Но и кроме того, что более важно, машина в высшей степени нечувствительна к *продолжающемуся* существованию яблок, полей, колоколен и т. д. Даже если бы все эти вещи *прекратили* существовать, машина все же успешно вела бы речь прежним образом. Вот почему вообще нельзя рассматривать машину как осуществляющую референцию.

Для нашей дискуссии важно заметить, что в тесте Тьюринга нет ничего, что отсеяло бы машину, запрограммированную лишь на имитационную игру и не умеющую делать *ничего* другого, а машина, не умеющая делать ничего, кроме как вести имитационную игру, *явно* указывает на что-либо не в большей степени, чем магнитофон.

МОЗГИ В БОЧКЕ (ЕЩЕ РАЗ)

Сравним гипотетические «мозги в бочке» с описанными машинами. Очевидны важные различия. У мозгов в бочке нет органов чувств, но есть *входы для данных* органов чувств; т. е. есть центростремительные нервные окончания, есть ввод данных с этих центростремительных нервных окончаний, и эти входные данные играют такую же роль в «программе» мозгов в бочке, что и в программе нашего мозга. В силу этих причин может показаться абсурдным отрицать их наделенность сознанием или ра-

зумом. Но тот факт, что они сознательны и разумны, не означает, что их слова указывают на то же, на что и наши. Интересующий нас вопрос таков: действительно ли их вербализации, содержащие, скажем, слово «дерево», указывают на *деревья*? В более общем виде: могут ли они вообще указывать на *внешние* объекты? (В противоположность, например, объектам, составляющим образ, порожденный автоматическим механизмом).

Для того, чтобы сформулировать наши идеи, уточним, что автоматические механизмы, как предполагается, появились в результате какой-то космической случайности или совпадения (или, может быть, всегда существовали). В этом гипотетическом мире у автоматических механизмов нет разумных создателей-проектировщиков. Фактически, как мы заметили в начале главы, мы можем представить, что все наделенные ощущениями существа (как минимальна бы ни была их чувствительность) находятся внутри бочки.

Это допущение не способствует достижению успеха. Нет связи между *словом* «дерево», употребляемым этими мозгами, и настоящими деревьями. Они будут употреблять слово «дерево» тем же способом, будут иметь те же мысли и образы, даже если деревьев не будет существовать. Их образы, слова и т. д. количественно идентичны с образами, словами и т. д., которые представляют деревья в *нашем* мире; но мы уже видели (вспомним муравья!), что качественное подобие чему-либо, что представляет объект (Уинстона Черчилля или дерево), само по себе не дает репрезентации. Коротко говоря, мозги в бочке не думают о реальных деревьях, когда они думают «передо мной дерево», потому что нет ничего такого, в силу чего их мысль «дерево» представляла бы реальное дерево.

Если такой вывод покажется поспешным, подумайте над следующим: мы видели, что слова не обязательно указывают на деревья, даже если они организованы в последовательность, идентичную такой речи, которая (если бы она возникла в одном из наших сознаний) безусловно *относилась бы к деревьям в действительном мире*. Так же и «программа» — в смысле правил, практик, диспозиций мозгов к вербальному поведению — не указывает с необходимостью на деревья и не вызывает референцию к деревьям через те связи, которые она устанавливает между словами и другими словами, между *языковыми* репликами и *языковыми*

ответами. Если эти мозги думают о деревьях, имеют референцию к деревьям, репрезентируют деревья (реальные деревья, находящиеся за пределами бочки), то это возможно в силу того, что «программа» соединяет языковую систему с *невербальными* входом и выходами. В мире Мозгов-в-Бочке существуют такие невербальные входы и выходы (опять эти центробежные и центростремительные нервные окончания!), но мы уже видели, что «чувственные данные», произведенные автоматическими механизмами, не репрезентируют деревья (или что бы то ни было внешнее), даже если они в точности схожи с нашими образами деревьев. Подобно тому, как пятна краски могут походить на изображение дерева, *не являясь* изображением дерева, так же, мы видели, «чувственные данные» могут быть количественно идентичны «образу дерева», не будучи образом дерева. Как может тот факт, что, в случае мозгов в бочке, язык связан программой с сенсорными входами, которые не репрезентируют, ни внутренне, ни внешне, деревья (или что бы то ни было внешнее) — как может это привести к тому, что полная система репрезентаций, используемый язык, указывает на нечто внешнее или репрезентирует его?

Ответ состоит в том, что это невозможно. Полная система чувственных данных, двигательные сигналы на центробежных окончаниях и вербально или концептуально опосредованная мысль, связанная «правилами языкового входа» с чувственными данными (или чем бы то ни было) в качестве вводных данных и «правилами языкового выхода» с двигательными сигналами в качестве выходных данных — эта система не более связана с *деревьями*, чем муравьиная кривая с Уинстоном Черчиллем. Как только мы увидели, что *качественное подобие* (приводящее, если хотите, к качественной идентичности) между мыслями мозгов в бочке и чьими-то мыслями в реальном мире никоим образом не ведет к подобию референции, то становится ясно, что совершенно нет оснований рассматривать мозг в бочке как указывающий на внешние объекты.

ПОСЫЛКИ АРГУМЕНТА

Итак, я привел обещанный аргумент, который должен был показать, что мозги в бочке не могут подумать или сказать, что

они — мозги в бочке. Остается только дать его окончательную, явную формулировку и исследовать его структуру.

Исходя из того, что было сказано выше, мы принимаем, что когда мозг в бочке (в мире, где каждое наделенное ощущениями существо является и всегда было мозгом в бочке) думает: «Вот передо мной дерево», то его мысль не указывает на настоящие деревья. Согласно некоторым теориям, которые мы будем обсуждать, она может указывать на образы деревьев или на электронные импульсы, которые вызывают представления о деревьях, или на детали программы, управляющие этими импульсами. Тем, что было сказано выше, эти теории не исключаются, поскольку существует тесная каузальная связь между употреблением слова «дерево» в бочковом английском и наличием образов деревьев, наличием электронных импульсов и наличием определенных деталей машинной программы. Согласно этим теориям, мозг *прав*, а не *ошибается*, когда думает: «Вот передо мной дерево». Принимая во внимание то, на что в бочковом английском указывает слово «дерево», и то, на что указывает слово «перед»; принимая, далее, что одна из этих теорий правильна, мы получим, что условия истинности предложения «Вот передо мной дерево» в бочковом английском языке состоят в том, что образ дерева состоит с обсуждаемым «я» (принадлежащим той же образной системе) в отношении «быть перед». Или же, возможно, эти условия истинности состоят в том, что электронные импульсы того вида, который обычно производит такой опыт, исходят из автоматических механизмов; или же — в том, что задействованы те особенности или детали этих механизмов, которые, как предполагается, производят знание «Вот передо мной дерево». И эти условия истинности явно выполняются.

В силу того же аргумента «бочка» в бочковом английском указывает на образ бочки или что-либо связанное с ним (электронные импульсы или детали программы), но явно не на реальную бочку, поскольку употребление слова «бочка» в бочковом английском каузально не связано с реальными бочками (за исключением связи, состоящей в том, что мозги в бочке не могли бы употреблять слово «бочка», если бы не наличие конкретной бочки — той, в которой они сидят; но эта связь возникает и между употреблением *каждого* слова в бочковом английском и этой конкретной бочкой; это не особенная связь между употреблением

конкретного слова «бочка» и бочками). Точно так же «питательный раствор» указывает в бочковом английском на образ жидкости или что-либо связанное с ним (электронные импульсы или детали программы). Следовательно, если их «возможный мир» в действительности является реальным миром, а мы в действительности представляем собой мозги в бочке, тогда то, что мы теперь подразумеваем под выражением «Мы — мозги в бочке» — это что мы *образно* представляем собой мозги в бочке или что-то в этом роде (если мы вообще подразумеваем этим что бы то ни было). Но в гипотезу о том, что мы являемся мозгами в бочке, входит представление о том, что мы не являемся образом мозгов в бочке (т. е. то, что мы «галлюцинируем» — это не что мы мозги в бочке). Итак, если мы — мозги в бочке, то предложение «Мы — мозги в бочке» сообщает нечто ложное (если оно вообще сообщает что бы то ни было). Короче говоря, если мы — мозги в бочке, то «Мы — мозги в бочке» ложно. Следовательно, это (с необходимостью) ложно.

Предположение о том, что такая возможность имеет смысл, возникает из сочетания двух ошибок: (1) слишком серьезного отношения к *физической возможности* и (2) бессознательного обращения к магической теории референции — теории, согласно которой определенные ментальные репрезентации с необходимостью указывают на определенные внешние предметы и виды предметов.

Имеется «физически возможный мир», в котором мы являемся мозгами в бочке — что это значит, кроме того, что имеется *описание* такого положения дел, которое совместимо с законами физики? Поскольку в нашей культуре (начиная еще с XVII столетия) существует тенденция воспринимать *физику* как метафизику, т. е. рассматривать точные науки как долгожданное описание «полного и окончательного содержания вселенной», постольку существует, как ее неосредственное следствие, тенденция принимать «физическую возможность» за истинный критерий того, что может иметь место в действительности. С такой точки зрения, истина есть физическая истина; возможность — физическая возможность, а необходимость — физическая необходимость. Но мы только что видели, даже если пока лишь на весьма искусственном примере, что такие взгляды ошибочны. Существование «физически возможного мира», в кото-

ром мы являемся мозгами в бочке (всегда были ими и всегда будем) не означает, что мы можем в действительности, по-настоящему, с определенной вероятностью *быть* мозгами в бочке. И устраняет эту возможность не физика, а *философия*.

Некоторые философы, стремясь предъявить и одновременно свести к минимуму требования своей профессии (типичное состояние умов в англо-американской философии XX века), сказали бы: «Ну да, вы показали, что некоторые из тех вещей, которые кажутся физически возможными, в действительности являются *концептуально* невозможными. Что же в этом удивительного?»

Разумеется, мой аргумент может быть охарактеризован как «концептуальный». Но характеристика философской деятельности как поиска «концептуальных» истин превращает ее в *исследование значений слов*. А это вовсе не то, чем мы собирались заняться.

Наша задача — рассмотрение *предварительных условий мышления о чем-либо, репрезентации чего-либо, указания на что-либо* и т. п. Мы исследовали эти предпосылки *не* путем исследования значений этих слов и фраз (как это мог бы сделать, например, лингвист), а путем *априорного рассуждения* — не в старом «абсолютном» смысле (поскольку мы не заявляем, что магические теории референции *априорно* ошибочны), но в смысле исследования того, как возможно *разумно допустить* определенные общие посылки или в смысле принятия определенных, достаточно широких теоретических допущений. Такая процедура не является ни «эмпирической», ни вполне «априорной», но включает в себя элементы обоих способов исследования. Несмотря на погрешность моей процедуры и на ее зависимость от допущений, которые могут быть охарактеризованы как «эмпирические» (например, допущение о том, что сознание не имеет другого доступа к внешним вещам или свойствам, кроме предоставляемого чувствами), она тесно связана с тем, что Кант называл «трансцендентальным» исследованием, поскольку это — повторяю еще раз — исследование *предпосылок* референции и тем самым мышления — предпосылок, встроенных в саму природу нашего сознания, хотя и не полностью независимых (как надеялся Кант) от эмпирических допущений.

Одна из посылок аргумента очевидна и состоит в том, что

магические теории референции ошибочны — ошибочны не только для физических, но и для ментальных репрезентаций. Другая посылка состоит в том, что невозможно указывать на определенные виды вещей без всякого каузального взаимодействия с ними³ или с теми вещами, в терминах которых они могут быть описаны. Но почему мы должны принимать эти послышки? Поскольку они представляют широкий фон, на котором я рассуждаю, пора рассмотреть их подробнее.

ПРИЧИНЫ ОТРИЦАНИЯ НЕОБХОДИМЫХ СВЯЗЕЙ МЕЖДУ РЕПРЕЗЕНТАЦИЯМИ И ИХ РЕФЕРЕНТАМИ

Как я уже упоминал, некоторые философы (наиболее известный из которых Brentano) приписывали сознанию некоторую силу — «интенциональность», — в точности позволяющую ему осуществлять референцию. Очевидно, я отбросил эту идею как не дающую решения. Но что дает мне право так поступить? Может быть, я поспешил?

Эти философы не утверждали, что мы можем мыслить о внешних вещах или свойствах, вообще не используя репрезентаций. Приведенный мной при сравнении чувственных данных с муравьиной «картиной» аргумент (аргумент от научно-фантастической истории о «картине» дерева, возникшей из пролитой краски и породившей чувственные данные, качественно подобные нашим «визуальным образам деревьев», но не сопровождавшейся никаким *понятием* дерева) был бы принят ими как показывающий, что *образы* не осуществляют референцию с необходимостью. Если ментальные репрезентации осуществляют референцию (к внешним вещам) с необходимостью, то они должны отражать природу *понятий*, а не природу образов. Но что такое *понятия*?

При интроспекции мы не воспринимаем как таковые «понятия», проплывающие по сознанию. Остановив поток мысли во

3. Если бы Мозги в Бочке имели каузальную связь, скажем, с деревьями в будущем, то, возможно, они могли бы указывать на деревья *сейчас* при помощи дескрипции «вещи, на которые я буду указывать как на «деревья» в такой-то и такой-то момент в будущем». Но мы должны представить себе такой случай, в котором Мозги в Бочке *никогда* не выберутся из бочки, а следовательно, *никогда* не вступят в каузальную связь с деревьями и т. д.

времени или в пространстве. мы обнаружим слова, образы, ощущения, чувства. Когда я высказываю вслух свои мысли, я не мыслю их дважды. Я слышу свои слова по мере того, как говорю. Разумеется, я чувствую разницу между тем, когда я произношу слова, которым верю, и когда я произношу слова, которым не верю (но иногда, когда я нервничаю или перед враждебной аудиторией, я могу чувствовать себя как если бы я лгал, тогда, когда я говорю правду); и я чувствую разницу между тем, когда я произношу слова, которые понимаю, и когда я произношу слова, которые не понимаю. Но я без труда могу представить себе кого-либо, кто мыслит именно эти слова (т. е. проговаривает их в уме) и переживает то же самое понимание, утверждение и т. п., что и я, но через минуту (или будучи разбужен гипнотизером) осознает, что вообще не понимал того, что только что прошло через его сознание; что он даже не понимает того языка, которому принадлежат эти слова. Я не утверждаю, что это весьма вероятно: я просто хочу сказать, что в этом нет ничего невообразимого. И это показывает нам не то, что понятия *и есть* слова (или образы, ощущения и т. д.), а что приписывать кому-либо «понятие» или «мысль» — это совсем не то, что приписывать ему ментальные «представления», какие бы то ни было сущности или события, доступные внутреннему самонаблюдению. Понятия не являются ментальными представлениями, в силу присущих им свойств указывающими на внешние объекты — по той решающей причине, что они вообще не являются ментальными представлениями. Понятия — это употребленные определенным способом знаки; они могут быть публичными или частными, ментальными или физическими сущностями, но даже когда знаки являются «ментальными» и «частными», сам по себе знак, в отрыве от своего использования, еще не есть понятие. И знаки не обладают референцией в силу своих свойств.

Мы можем увидеть это, проделав очень простой мысленный эксперимент. Предположим, что вы, как я, не можете отличить вяза от бука. Мы все же скажем, что референция слова «вяз» в моей речи — та же, что и референция слова «вяз» в речи других людей, а именно деревья-вязы, и что множество всех букв является экстенсионалом термина «бук» (т. е. множеством всех вещей, которым может быть истинно предсказано слово «бук») — как в моей, так и в вашей речи. Можно ли согласиться,

что различие между тем, на что указывает «вяз» и тем, на что указывает «бук», вызвано различием в наших *понятиях*? Со стыдом признаю, что мое понятие вяза точно таково, как мое понятие бука. (Кстати, это показывает, что референция определена социально, а не индивидуально: и вы, и я считаемся с мнением экспертов, *умеющих* отличить бук от вяза.) Если же кто-то отважно возьмется утверждать, что различие между референцией «вяза» и референцией «бука» в *моей* речи объясняется различием моих психологических состояний, то пусть представит себе Землю-Двойник, где эти слова поменялись местами. Земля-Двойник очень похожа на Землю; по сути, за исключением того факта, что «вяз» и «бук» заменены друг на друга, читатель может предположить, что Земля-Двойник похожа на Землю в точности. Предположим, что на Земле-Двойнике у меня есть личный двойник, идентичный со мной до молекулы (в том смысле, в котором могут быть «идентичными» два галстука). Если вы — дуалист, предположите, что мой двойник мыслит те же вербализованные мысли, что и я, обладает теми же чувственными данными, теми же диспозициями и т. д. Абсурдно было бы думать, что его психологическое состояние отлично от моего, однако же его слово «вяз» репрезентирует *буки*, а мое слово «бук» репрезентирует *вяза*. (Точно так же, если «вода» на Земле-Двойнике — это какая-то другая жидкость, скажем, XYZ, а не H₂O, то «вода» будет репрезентировать разные жидкости на Земле-Двойнике и на Земле и т. д.) Вопреки доктрине, сопутствующей нам с XVII века, *значения не находятся в голове*.

Мы видели, что обладание понятием — это не вопрос обладания образами (например, деревьев — или даже образами, «визуальными» или «акустическими», предложений или целых дискурсов), так как можно обладать какой угодно системой образов и не иметь способности употреблять предложения ситуационно подходящими способами (учитывая в определении «ситуационной уместности» как лингвистические факторы — о чем говорилось ранее, — так и внелингвистические). Человек может обладать всеми образами, какими угодно, но тем не менее испытывать сильное затруднение, когда его попросят «указать на дерево» даже в присутствии множества деревьев. Он может даже обладать образом того, что от него ожидается, и все же не знать, чего от него ожидают. Ведь если образ не сопро-

вождается способностью действовать определенным образом, то он — всего лишь *картинка*, а действовать в соответствии с картинкой — это уже способность, которой можно владеть или не владеть. (Человек может изобразить себя указывающим на дерево лишь ради созерцания чего-то логически возможного; указывающим на дерево после того, как некто произнес (для него бессмысленную) последовательность звуков «укажите на дерево»). Он все еще не будет знать, что от него ожидалось указание на дерево, и все еще не будет *понимать* выражение «укажите на дерево».

Я рассмотрел способность употреблять определенные выражения как критерий обладания полноценным понятием, но этот критерий легко может быть ослаблен. Мы можем допустить, например, такую символику, которая состоит из элементов, не являющихся словами естественного языка, и такие ментальные феномены как образы и другие типы внутренних событий. Здесь существенно то, что они должны обладать такой же сложностью, такой же способностью сочетаться друг с другом и т. д., что и предложения естественного языка. Хотя для некоторого математика определенное представление — например, синяя вспышка — может служить внутренним выражением целого доказательства теоремы простых чисел, но вряд ли нам придет в голову утверждать (и было бы ложным утверждать) это, если бы этот математик не мог развсрнуть свою «синюю вспышку» в отдельные шаги и логические связи доказательства. Но, вне зависимости от того, какой вид внутренних феноменов мы допускаем в качестве возможных *выражений* мысли, аргументы, в точности подобные приведенному выше, показывают, что понимание конституируется вовсе не феноменами самими по себе, но способностью мыслителя *применять* эти феномены, производить правильные феномены в правильных обстоятельствах.

Вышеизложенное представляет собой весьма сокращенный вариант аргумента, данного Витгенштейном в «Философских исследованиях». Если он справедлив, то попытка понять мышление с помощью того, что называют «феноменологическим» исследованием, исходит из неверных оснований, поскольку феноменологи не видят, что описывают внутреннее *выражение* мысли, тогда как *понимание* этого выражения — чье-либо понимание собственных мыслей — представляет собой не *явление или*

событие, но способность. Наш пример с человеком, якобы мыслящим по-японски (и обманувшим японского телепата), уже показывает тщетность феноменологического подхода к проблеме понимания. Даже если есть какое-то доступное интроспекции свойство, присутствующее тогда и только тогда, когда некто действительно понимает нечто (хотя интроспективно это кажется ложным), все же это свойство лишь *коррелирует* с пониманием, и возможно, что человек, обманывающий японского телепата, также обладает этим качеством, но *в то же время* не понимает ни слова по-японски.

С другой стороны, рассмотрим вполне возможного человека без какого бы то ни было «внутреннего монолога». Он говорит на прекрасном английском и на просьбу высказать свое мнение по тому или иному вопросу даст пространный ответ. Но он никогда не мыслит (в словах, образах и т. д.), когда не высказывается вслух; ничто не «проходит через его голову» — за исключением, конечно, того, что он слышит свой собственный голос и располагает обычными чувственными впечатлениями о своем окружении, а также общим «чувством понимания». (Может быть, у него есть привычка разговаривать с самим собой.) Когда он печатает письмо, идет в магазин и т. д., у него нет никакого внутреннего «потока мысли»; однако его поступки разумны и целесообразны, и если кто-то подойдет и спросит его «Что вы делаете?», он даст вполне связный ответ.

Такого человека вполне можно себе представить. Мы не колеблясь, станем утверждать, что он обладает сознанием, что он, — если он неоднократно выражал стойкое отвращение к рок-н-роллу, — не любит рок-н-ролл и т. д., лишь на том основании, что у него нет осознанных мыслей, за исключением тех случаев, когда он высказывается вслух.

Из этого следует, что (а) никакое множество ментальных событий — образов или более «абстрактных» ментальных происшествий и свойств — *не конституирует* понимание; и что (б) никакое множество ментальных событий не является *необходимым* для понимания. В частности, *понятия не могут быть идентичны ментальным объектам* любого вида. Мы видели, что — в предположении, что под ментальным объектом мы подразумеваем нечто доступное интроспекции — чем бы ни являлся такой объект, его может не быть у человека, который понимает соответствующее

слово (и тем самым обладает полноценным понятием), и он может наличествовать у человека, который вообще не обладает этим понятием.

Возвращаясь к нашей критике магических теорий референции (тема, также связанная с Витгенштейном), мы видим, что, с одной стороны, те «ментальные объекты», которые мы *можем* обнаружить интроспективно — слова, образы, чувства и т. д. — не указывают в силу своих свойств на что бы то ни было в большей степени, нежели это делает муравьиная картинка (и это происходит по тем же причинам). В то же время попытки постулировать особые ментальные объекты, «понятия», которые *обладают* необходимой связью со своими референтами и которые может обнаружить лишь подготовленный феноменолог, представляют собой логическую ошибку, поскольку понятия являются (по крайней мере частично) *способностями или умениями*, а не явлениями. Доктрина, согласно которой ментальные репрезентации могут с необходимостью указывать на внешние вещи, плоха не только с естественнонаучной точки зрения — она плоха также феноменологически и ведет к путанице в понятиях.

2. ПРОБЛЕМА РЕФЕРЕНЦИИ

Почему нас так удивляет, что гипотеза о Мозгах в Бочке оказывается внутренне противоречивой? Причина этого в том, что мы склонны думать, будто то, что *происходит в наших головах*, должно определять то, что мы подразумеваем и на что указывают наши слова. Но нетрудно увидеть, что это не так. Тривиальный контрпример дают обычные индексикальные слова — такие, как *я, это, здесь, сейчас*. Когда я думаю, например, «Я опаздываю на работу», я могу находиться в том же ментальном состоянии, что и Генри (представим себе, что мы с Генри полностью идентичные близнецы), но все же токен слова «я», возникающий в моем мышлении, указывает на меня, а токен слова «я», возникающий в мышлении Генри, указывает на *Генри*. Я могу находиться в одном и том же ментальном состоянии¹, когда я думаю «Я опаздываю на работу» во вторник и когда я думаю «Я опаздываю на работу» в среду, но время, на которое указывает временная форма глагола «опаздываю», различно в этих двух случаях. Более изощренный пример этого положения — применение терминов естественных видов.

Предположим, развивая упомянутый в предыдущей главе случай, что на Земле-Двойнике люди говорят по-английски (каким-то чудесным образом они развивались так, что в точности походят на нас и говорят на языке, который — за одним исключе-

1. По крайней мере, я могу быть «в том же ментальном состоянии» в том отношении, что психологический процесс, результатом которого в моем мышлении становится эта мысль, может иметь параметры с теми же значениями. Мое *общее* ментальное состояние будет, без сомнения, другим, поскольку во вторник я считаю, что сегодня — вторник, а в среду я так не считаю. Однако согласно той теории, в которой значения слов зависят от моего *общего* ментального состояния и изменяются при каждом его изменении, никакие слова никогда не могли бы иметь одинаковые значения; такая теория привела бы к отказу от самого понятия значения слова. Более того, мы могли бы придумать такую историю из жизни на Земле-Двойнике, в которой у нас с моим двойником одинаковые *общие* ментальные состояния, но референция слов «я» и «сейчас» все же различна (поскольку календарь на Земле-Двойнике не синхронизован с нашим).

нием, о котором я сейчас скажу — идентичен английскому, каким он был лет двести назад). Я буду исходить из того, что эти люди еще не знают дальтоновой или пост-дальтоновой химии. В частности, они не владеют такими понятиями, как « H_2O ». Предположим, что реки и озера на Земле-Двойнике наполнены жидкостью, которая поверхностно напоминает воду, но которая *не* является H_2O . Тогда слово «вода», употребленное на Земле-Двойнике, указывает *не* на воду, а на эту другую жидкость (например, XYZ). Однако нет никаких релевантных различий в ментальных состояниях людей, употребляющих язык на Земле-Двойнике, и людей, употребляющих язык на Земле (скажем, в 1750 году), которые могли бы объяснить это различие в референции. Референция отличается потому, что отличается *вещество*². Само по себе ментальное состояние, в отрыве от ситуации в целом, не устанавливает референцию.

Однако некоторые философы критикуют этот пример, утверждая, что, если бы такая планета была открыта, то мы говорили бы, что «есть два вида воды», а *не* что наше слово «вода» не указывает на жидкость Земли-Двойника. Согласно этим критикам, если бы мы обнаружили реки и озера, полные похуже на воду жидкости, которая при том — не H_2O , то мы тем самым фальсифицировали бы утверждение о том, что вся вода — H_2O .

Легко усовершенствовать наш пример так, чтобы избежать этого аргумента. Прежде всего, жидкость Земли-Двойника вовсе не обязательно должна быть *настолько* похожа на H_2O . Предположим, что на самом деле это смесь, состоящая из 20% хлебного спирта и 80% воды, а химия тел жителей Земли-Двойника такова, что они не только не отравляются, но и вообще не чувствуют никакой разницы между такой смесью и H_2O . Такая жидкость будет отличаться от воды во многих отношениях; однако *типичный* носитель языка может не знать об этих различиях, а потому будет в точности в том же ментальном состоянии, что и *типичный* носитель языка на Земле в 1750 году. Конечно, «вода» Земли-Двойника по вкусу отличается от земной воды *для нас*, но она не отличается по вкусу *для них*. Она иначе ведет себя при

2. Более развернутое обсуждение этой темы см. в статье «Значение «значения»» (Патнэм Х. Философия сознания / Пер. с англ. М., ДИК, 1999. С. 164–234).

кипячении, но должен ли говорящий по-английски знать, когда в точности вода кипит и что в точности при этом происходит, для того, чтобы ассоциировать со словом «вода» некоторое совершенно стандартное понятийное содержание?

Можно возразить, что вполне могут существовать такие эксперты по Земле-Двойнику, которые знают о «воде» такие вещи (например, что это — *смесь* двух жидкостей), которых мы не знаем о воде (в которых мы не были убеждены относительно воды в 1750 году, потому что они не *истинны* относительно воды), и, следовательно, что *коллективное* ментальное состояние говорящих по-английски на Земле-Двойнике отличается от коллективного ментального состояния людей, говорящих по-английски на Земле (в 1750 году). Можно признать, что референция чьего-либо термина не устанавливается индивидуальным ментальным состоянием говорящего, и при этом настаивать, что референцию термина устанавливает полное ментальное состояние *всех* членов языкового сообщества.

Трудность здесь состоит в том, что люди на Земле *или* Земле-Двойнике могли бы и не развить химию к 1750 году до такой степени. Если термин в своем общепринятом употреблении имел на Земле до развития химии то же самое значение и референцию, что и сегодня, и если термин имел то же самое значение и референцию на Земле-Двойнике до развития соответствующего знания, что и сегодня, то мы можем вернуться к тому более раннему времени, когда коллективные ментальные состояния этих двух сообществ были *одинаковы* во всех отношениях, релевантных для назначения экстенционала термина «вода», и утверждать, что экстенционал был иным *тогда* (как и теперь) и что *коллективное* ментальное состояние не устанавливает экстенционал. Скажем ли мы тогда, что референция *изменилась* с развитием химии? Что термин *раньше* указывал на оба вида воды (несмотря на разность во *вкусе* для нас!) и стал указывать на *различные* виды только после того, как до определенной степени развилась химия?

Если мы говорим, что референция их терминов или наших терминов изменилась, когда они или мы развили химию (вплоть до способности дистиллировать жидкости, различать, что вода со спиртом — это смесь и т. д.), то мы должны будем сказать, что почти каждое научное открытие изменяет референцию наших

терминов. Мы не *обнаруживали*, что вода (в донаучном смысле) является H_2O согласно такому представлению; скорее мы *постулировали* это. Мне это кажется явно неправильным. Мы постоянно полагали водой что бы то ни было, что имело ту же природу, что и местное вещество, отличающееся при помощи этого термина; и мы обнаружили, что вода *в этом смысле* была H_2O . Люди же Земли-Двойника постоянно полагали «водой» некоторое вещество в их окружающей среде, отличающееся при помощи этого термина, и их эксперты обнаружили, что «вода» *в этом смысле* — смесь двух жидкостей.

Если мы соглашаемся, что «вода» не изменяет *значение* (в каждом из языков), когда эксперты делают такие открытия, как «вода — это H_2O » или «вода — это смесь двух жидкостей», или не меняет свое обычное значение и референцию (конечно, в результате таких открытий могут появиться более технические способы использования термина), и что «вода» в обычном земном значении и при обычной земной референции не включает смеси алкоголя и воды, то мы должны сказать, что опытное знание не играет никакой роли в объяснении разницы в *значении* слова «вода» на Земле и на Земле-Двойнике. И при этом оно не влияет на объяснение *референции*, поскольку мы могли бы рассмотреть еще одну Землю-Двойник, где вода была бы *другой* смесью, а опытное знание было бы тем же самым (довольно скудным) опытным знанием, что и на первой Земле-Двойнике. Или, как было только что замечено, мы могли бы просто вообразить, что эксперты по Земле и по Земле-Двойнику еще не существуют. Слово «вода» тогда все еще *указывало бы на другое вещество*, даже если *коллективное* ментальное состояние в этих двух сообществах было бы одним и тем же. То, что происходит в головах людей, не устанавливает референцию их терминов. Работу по установлению экстенционала термина завершает, по выражению Милля, «сама субстанция».

Когда мы видим, что ментальное состояние (и в индивидуальном, и в коллективном смысле) не устанавливает референцию, то нас уже не удивляет, что Мозги в Бочке не могут преуспеть в указании на внешние объекты (даже при том, что их ментальные состояния те же, что и у нас), и, следовательно, не могут говорить или думать, что они — Мозги в Бочке.

НАМЕРЕНИЯ. ЭКСТЕНСИОНАЛЫ И «ПОНЯТИЙНЫЕ МИРЫ».

Чтобы рассмотреть проблему того, как устанавливается референция наших терминов (при том, что она не устанавливается просто нашими ментальными состояниями), нам будет удобно иметь в распоряжении некоторые технические термины. В логике множество тех вещей, относительно которых термин истинен, называется *экстенсионалом* этого термина. Таким образом, экстенсионал термина «кошка» — множество кошек. Если термин имеет больше чем один смысл, то мы притворяемся, что у этого слова есть невидимые приписки (так, что это на самом деле как бы два слова, а не одно), например, «кролик₁» — экстенсионал: множество кроликов; «кролик₂» — экстенсионал: множество трусов. (Строго говоря, экстенсионал термина естественного языка всегда несколько нечеток, но мы притворимся для простоты, что промежуточные случаи так или иначе подчиняются правилам.)

Такие слова, как «я», которые указывают в разных случаях на разных людей, будут иметь не экстенсионал, а *экстенсиональную функцию*: это функция, которая определяет экстенсионал в каждом контексте использования. В случае слова «я» экстенсиональная функция довольно проста; это просто функция $f(x)$, значением которой для любого говорящего x является множество, состоящее только из x . Аргумент « x » определяется по релевантному параметру, используемому для описания контекста (в данном случае, по *говорящему*), и упоминается в семантике как *индекс* (*показатель*). Индексы необходимы для обозначения времен, объектов демонстративного указания и других деталей контекста при полной семантической трактовке, но опустим детали.

Множество вещей, которое составляет экстенсионал термина «кошка», будет разным в разных возможных ситуациях или «возможных мирах». В возможном мире M , в котором нет никаких кошек, экстенсионал термина «кошка» является пустым множеством. Если бы у моей кошки Эльзы было потомство, то экстенсионал термина «кошка» имел бы по крайней мере один член, которого он не имеет в фактическом мире. (Мы можем выразить это, сказав, что в каждом возможном мире M , в котором у Эльзы есть потомство, экстенсионал термина «кошка»

включает таких членов, которых он не включает в действительном мире).

Мы можем показать то отношение, в котором экстенционал термина изменяется вместе с возможным миром M , точно тем же способом, которым мы показываем, как экстенционал слова «я» изменяется вместе с говорящим: используя *функцию*. Мы принимаем, что множество абстрактных объектов, называемых «возможными мирами», представляет различные состояния дел или возможных мировых историй, и связываем с термином «кошка» функцию $f(M)$, значение которой в каждом возможном мире M – множество возможных объектов, которые являются кошками в мире M . Эту функцию, вслед за Монтегью и Карнапом, я буду называть *интенсионалом*³ слова «кошка». Точно же интенционал двуместного предиката «касается» является функцией $f(M)$, значение которой в любом возможном мире M – множество упорядоченных пар возможных объектов, которые касаются друг друга в мире M . Интенционал трехместного предиката « x находится между y и z » – функция, значением которой в любом возможном мире является множество таких упорядоченных троек (x, y, z) , что x расположен между y и z , и так далее. Интенсионалом таких слов, как «я», экстенционал которых в любом мире зависит от контекста, будет более сложная функция, имеющая своими аргументами и возможный мир, и индексы, представляющие контекст.

Интенционал уточняет то, как экстенционал зависит от возможного мира. Таким образом, он очень полно представляет то, что нас интересует – связанный с термином экстенционал, так как он сообщает нам, чем этот экстенционал был бы в любом возможном мире.

Причина, по которой «интенционал» (в этом смысле) не может быть отождествлен со *значением*, состоит в том, что любые два логически эквивалентных термина имеют один и тот же экстенционал в каждом возможном мире и, следовательно, один и тот же интенционал. Однако теория, которая не может провести различие между терминами с одинаковыми значениями

3. Montague R. Formal Philosophy. Yale University, 1974. Это использование термина «интенционал» – не то традиционное использование, которое я обсуждал в статье «Значение «значения»».

и терминами, которые эквивалентны лишь в логике и математике, представляет собой неадекватную теорию значения. «Куб» и «правильный многогранник с шестью квадратными сторонами» суть логически эквивалентные предикаты. У этих двух терминов одинаковый интенционал, а именно функция, значение которой в любом возможном мире — множество кубов в этом мире; но есть различие в *значении*, которое было бы потеряно, если бы мы просто отождествили значение с этой функцией.

Позвольте мне подчеркнуть, что в этой теории нам следует думать о возможных мирах, множествах и функциях как об абстрактных внементальных сущностях, а не путать их с представлениями или описаниями этих объектов.

Фреге считал, что значение (*Sinn*) выражения — некая внешняя по отношению к сознанию сущность или идея, которая может быть так или иначе «схвачена» сознанием. Такая теория будет бесполезна для интенционалов в *нашем* смысле. Во-первых, как было только что отмечено, есть различия в значении, которые не ухватываются интенционально; поэтому понимание термина не может состоять *только* в его соединении с интенционалом. Еще более важно то, что коль скоро мы предполагаем, что не имеем никакого «шестого чувства», которое позволяло бы нам *непосредственно* воспринимать внементальные сущности или каким-то образом иметь с ними дело (например, «интуитивно постигать» их), то «схватывание» интенционала или любой внементальной сущности должно быть некоторым способом опосредовано репрезентациями. (Это также кажется ясным интроспективно, по крайней мере мне). Но в целом проблема, которую мы исследуем, состоит в том, как репрезентации *могут* позволять нам указывать на то, что находится вне сознания. Принять понятие «схватывания» *X*, внешнего по отношению к сознанию, означало бы уклониться от решения вопроса.

Если я говорю о ком-то, что он «считает, что на столе стоит стакан воды», то я обычно приписываю ему способность указать на *воду*. Но, как мы видели, способность указать на воду требует прямой или косвенной связи с настоящей водой (H_2O); утверждение «Джон считает, что на столе перед ним стоит стакан воды» — не только утверждение о том, что происходит в голове Джона, но частично и утверждение о среде, окружающей Джона, и об отношениях, в которых Джон находится к этой окру-

жающей среде. Если оказывается, что Джон — житель Земли-Двойника, то, когда он говорит «на столе стоит стакан воды», он считает, что на столе стоит стакан, содержащий жидкость, которая *фактически* состоит из воды и хлебного спирта.

Гуссерль предложил полезный технический прием для описания того, что происходит в чьей-либо голове, без привлечения каких бы то ни было предположений о существовании или природе тех фактических вещей, на которые указывают наши мысли: это прием *заклучения в скобки*⁴. Если мы «закключаем в скобки» убеждение, которое мы приписываем Джону, когда говорим «Джон считает, что на столе стоит стакан воды», тогда то, что мы приписываем Джону, — это просто *ментальное состояние* действительного или возможного человека, который верит, что на столе стоит стакан воды (в полном, «не заключенном в скобки» обычном смысле). Таким образом, если Джон на Земле-Двойнике не может почувствовать разницу между водой и смесью воды и спирта, то он может находиться в том же ментальном состоянии, что и действительное или возможное существо, говорящее на земном английском языке, когда он говорит «на столе стоит стакан воды», несмотря на тот факт, что он имеет водой солидную выпивку. Мы будем говорить, что он имеет *заключенное в скобки убеждение* в том, что [на столе стоит стакан воды]. В действительности, принцип заключения в скобки вычитает из обычного выражения убеждения ряд следствий — все те отношения следования, которые указывают на внешний мир или на то, что является внешним по отношению к сознанию мыслящего.

Дэниел Деннетт недавно использовал выражение «понятийный мир» в смысле, связанном с тем, в котором Гуссерль использовал заключение в скобки⁵. Вся совокупность заключенных в скобки убеждений мыслящего составляет описание понятийного мира мыслящего, в смысле Деннетта. Таким образом, у людей на Земле-Двойнике примерно тот же самый *понятийный мир* (у них даже та же самая *понятийная вода*), что и у нас; однако они живут на другой *реальной планете* (и указывают на другое *действи-*

4. Husserl E. Ideas: General Introduction to Pure Phenomenology. L.: Allen & Unwin, 1969. (Впервые было опубликовано в 1913 году).

5. Dennett D. Beyond Belief // Thought and Object / Ed. Woodfield A. Oxford University Press, в печати.

тельное вещество как на «воду»). У Мозгов в Бочке из предыдущей главы мог бы быть тот же самый понятийный мир, что и у нас, до последней детали, если хотите; дело лишь в том, что ни один из их терминов не имел бы вообще никакой референции ко внешнему миру. Традиционная теория значения принимает, что понятийный мир мыслителя определяет интенционалы его терминов (и они, вместе с тем фактом, что некоторый возможный мир *М* является действительным, определяют экстенционалы терминов и истинностные значения всех предложений). Мы видели, что традиционная теория значения неправильна; и в этом причина того, почему литература сегодня содержит много различных концепций значения (например, «интенционал» и «понятийный мир»), а не единую общую концепцию. «Значение» *распалось на части*, и перед нами стоит задача собрать их. Если интенционал и экстенционал непосредственно не установлены понятийным миром, то как они установлены?

ПРИНЯТОЕ ПРЕДСТАВЛЕНИЕ ОБ ИНТЕРПРЕТАЦИИ

Наиболее общее представление о том, как нами устанавливаются интерпретации нашего языка — если не индивидуально, то коллективно — связан с понятиями *операциональных ограничений* и *теоретических ограничений*. Операциональные ограничения были первоначально задуманы довольно наивно; мы просто *постулируем* (так сказать, конвенционально), что некоторое предложение (скажем, «Электричество течет через этот провод») должно быть истинным, если и только если наблюдается некоторый результат испытаний (отклоняемая игла вольтметра, или, в феноменалистской версии, *наличие у меня визуального впечатления от наблюдения отклоняемой иглы вольтметра*). У грубого операционализма такого рода больше нет защитников, потому что было принято во внимание, что (1) связи между теорией и опытом являются вероятностными и не могут быть корректно формализованы как точные корреляции (даже если через провод идет ток, всегда могут иметь место события низкой вероятности или фоновые условия, которые удержат иглу вольтметра от отклонения); и (2) даже эти вероятностные связи не являются простыми *семантическими* корреляциями, но зависят от эмпирической теории, которая подвержена пересмотру. Согласно наи-

вному операционализму, значение и референция термина «золото» подвергались бы изменению каждый раз, когда обнаруживается новый способ проверки, является ли это вещество действительно золотом. (Фактически, мы не можем сказать, что *обнаруживается* новый способ проверки на золото). Согласно операционалистской картине, теории проверяются *предложение за предложением* (постулированные операциональные значения индивидуальных предложений сообщают, как проводить пошаговую проверку теории); согласно более современной картине, теория «проходит проверку опытом как единое целое», как выразился Куайн.

Можно, однако, ослабить понятие операционального ограничения таким образом, чтобы преодолеть или обойти все эти возражения. Так, можно ограничить класс интерпретаций (назначений интенционалов предикатам языка), которые будут приняты как допустимые, принципами следующей формы: «Допустимой является такая интерпретация, что *чаще всего* предложение *S* является истинным, когда выполнено опытное условие *E*» (соответственно, «такая, что *чаще всего* предложение *S* является *ложным*, когда *E* выполнено»). Такие ограничения формируют идею о том, что существуют *вероятностные* отношения между истинностью или ложностью языковых предложений и опытом. И, во-вторых, можно принять представление, согласно которому эти ограничения будут подвержены пересмотру по мере того, как развивается теория. Можно думать об этих ограничениях не как о постулировании значений, как это происходит в грубой версии операционализма, а как о пробных ограничениях, налагаемых на класс допустимых интерпретаций. Вслед за Пирсом (писавшем об этом за 50 лет до того, как Бриджмен заявил об «операционализме»!) можно принять представление, согласно которому идеальное множество операциональных ограничений само является чем-то таким, к чему мы последовательно приближаемся в ходе эмпирического исследования, а не чем-то таким, что мы лишь постулируем. Короче говоря, можно принять представление, согласно которому операциональные ограничения — это те, что *были бы* наложены рациональными исследователями, если бы последние наблюдали, экспериментировали и рассуждали столько, сколько возможно; те ограничения, которые они приняли бы в состоянии «рефлексивного равно-

весия», выбирая единственным образом интерпретацию наших терминов; те ограничения, которые мы фактически принимаем в любое данное время, имеют статус рациональной оценки или аппроксимации.

Такое представление совместимо с настойчивым утверждением Куайна о том, что связи между опытом и теорией настолько же подчинены пересмотру, насколько и любой другой аспект общего корпуса нашего знания. Не каждый такой пересмотр приводит к «изменению значения»: такие пересмотры могут быть (и часто являются) просто более успешными попытками определить то, о чем мы уже говорили, равно как и то, что прежде теория с операциональными ограничениями ухватывала лишь неадекватно.

В дополнение к ограничению класса допустимых интерпретаций посредством операциональных ограничений (или последовательных приближений к Пирсову идеальному множеству операциональных ограничений) можно также ввести ограничения, относящиеся к формальным свойствам теории. Например, «допустимая интерпретация такова, что при ней оказывается истинным, что различные следствия всегда имеют различные причины». Кант считал, что подобные «теоретические ограничения» являются непосредственно частью рациональности: мы скорее *налагаем* принцип детерминизма на мир, чем обнаруживаем его. В такой форме ограничение, конечно, слишком сильно: *ценой* сохранения детерминизма могло бы стать слишком большое усложнение нашей системы знания в целом. Но этот вид ограничения может быть смягчен так же, как и операциональные ограничения. (Например, можно требовать, чтобы детерминизм был сохранен всякий раз, когда «стоимость» этого в терминах усложнения других частей теории не слишком велика; в такой форме, как представляется, мы принимаем это ограничение.)

Теоретические ограничения часто бывают заявлены скорее как ограничения на принятие теории, чем как ограничения на интерпретацию теории, но им можно легко дать такое толкование, чтобы заставить их играть эту последнюю роль. Так, если автор заявляет ограничение «консервативизма» или «презервационизма» как ограничение на принятие теории («теория неприемлема, если она требует отказа от очень многих предвари-

тельно принятых убеждений и если доступна другая теория, настолько же «простая» в прочих отношениях, которая сохраняет эти убеждения и при этом согласуется с наблюдением»), то мы можем заново сформулировать это ограничение как ограничение на интерпретацию таким образом: «допустимая интерпретация такова, что дает истинные предложения, которые были приняты в течение долгого времени, за исключением тех случаев, где это потребует неуместного усложнения теории, состоящей из множества предложений, истинных при данной интерпретации, или слишком сильного пересмотра операциональных ограничений». Здесь также было широко распространено мнение, что невозможна никакая индуктивная логика, если мы не налагаем некоторое *априорное* упорядочение (называемое «упорядочением простоты» или «упорядочением правдоподобия») на гипотезы, которые могут быть приняты, имея определенные данные наблюдения (хотя упорядочение может само быть различным в различных контекстах эксперимента или наблюдения); ограничение, состоящее в том, что «множество предложений, которые являются истинными при допустимой интерпретации, не должно быть ниже по степени простоты, чем любое другое множество с теми же последствиями, известными нам благодаря наблюдению и опыту», соответствовало бы ограничению в индуктивной логике, состоящему в том, что следует принять наиболее простую (или наиболее «вероятную») из гипотез, совместных с наблюдениями.

Теоретические ограничения многих других видов были предложены в литературе по философии науки. Есть ограничения, которые, подобно «простоте», относятся к свойствам множества принятых предложений, и есть ограничения, которые относятся к *истории* исследования, в соответствии с которым это множество было принято. Но детали здесь нас не касаются. Причины приверженности идее о том, что допустимые интерпретации нашего языка (в смысле допустимых назначений интенсивных терминам языка) регулируются операциональными и теоретическими ограничениями, очевидны: есть ли у нас опыт некоторого вида — это вопрос, о котором сознание может судить (несмотря на философские проблемы «опыта»). Так, если теория подразумевает или содержит предложение, которое связано с опытом *Е* операциональным ограничением некоторого

вида, вероятностным или каким бы то ни было, то субъект может *знать*, работает ли теория, или же она не вполне пригодна, по крайней мере в этом случае, исходя из того, действительно ли он имеет опыт *Е*. Поскольку ограничения, которые мы используем, чтобы проверить теорию, *также* устанавливают экстенсионалы терминов, оценка субъектом «работы» теории – это в то же время оценка ее *истинности*. А поскольку знание говорящим этих ограничений – это знание интенсиналов терминов, схватывание правильной семантики сообщило бы нам, для любой предложенной теории мира, каков должен быть наш мир, чтобы теория была истинной.

Кроме того, если мы осуществляем процедуру идеализации, предполагая, что субъекты обладают тем, что экономисты называют «совершенной информацией» друг о друге, то каждый из них знает формальную структуру принятой теории *T* и прошлую историю программы исследования, которой она принадлежит, предыдущие убеждения, которые она сохраняет или нет, и т. д. Поэтому каждый мыслитель имеет возможность *знать*, выполнены теоретические ограничения или нет. (Если мы не хотим прибегать к идеализации, предполагающей совершенную информацию, то мы можем сказать, что это известно коллективу мыслящих).

Короче говоря, если бы принятое представление было правильно, то у нас была бы изящная теория того, как устанавливаются интенсиналы и экстенсиналы (в принципе – конечно, детали слишком сложны, чтобы мы могли проследить их на существующей стадии методологического знания). Но, к сожалению, принятое представление не работает!

ПОЧЕМУ ПРИНЯТОЕ ПРЕДСТАВЛЕНИЕ НЕ РАБОТАЕТ

Трудность с принятым представлением состоит в том, что оно содержит попытку устанавливать интенсиналы и экстенсиналы индивидуальных терминов, устанавливая условия истины для целых предложений. Идея, как мы только что видели, состоит в том, что операциональные и теоретические ограничения (те, которые рациональные исследователи приняли бы в некотором идеальном пределе исследования) определяют, какие предложения языка *истинны*. Однако даже если это правильно, такие ог-

раничения не могут определять то, на что наши термины *указывают*, так как в понятии операционального или теоретического ограничения нет ничего, что могло бы делать это непосредственно. Попытка же совершить указание *опосредованно*, вводя ограничения, которые выбирают множество истинных предложений, а затем надеясь, что, определяя истинностное значение целого предложения, мы можем так или иначе установить то, на что указывают *термины*, встречающиеся в этих предложениях, не достигает цели.

То, что такая попытка не достигает цели, показано Куайном⁶. Я хочу очень сильно расширить полученные им результаты о «неопределенности». Я буду доказывать, что, даже если у нас есть ограничения любого характера, которые определяют истинностное значение каждого предложения на некотором языке *в каждом возможном мире*, тем не менее, референция индивидуальных терминов остается неопределенной. Фактически, возможно интерпретировать полную систему языка совершенно различными способами, каждый из которых будет совместим с требованием того, что истинностное значение каждого предложения в каждом возможном мире должно быть определено. Короче говоря, мало того, что принятое представление не работает; *никакое представление, которое только устанавливает истинностные значения целых предложений, не может устанавливать референцию*, даже если оно определяет истинностные значения предложений *в каждом возможном мире*.

Подробное доказательство носит технический характер, и я думаю, что будет уместно привести его в приложении. Здесь же я дам только иллюстрацию метода доказательства, а не детальное доказательство.

Рассмотрим предложение:

- (1) Кошка [находится] на коврике. (Здесь и далее «находится» употреблено во *вневременном* смысле, то есть это означает «находилась, находится или будет находиться».)

В стандартной интерпретации это истинно в тех возможных мирах, в которых есть по крайней мере одна кошка на по край-

6. См.: Quine W. V. Ontological Relativity // Quine W. V. Ontological Relativity and Other Essays. Columbia University Press, 1969.

ней мере одном коврикe в некоторое время — прошлое, настоящее или будущее. Кроме того, «кошка» указывает на кошек, а «коврик» указывает на коврики. Я покажу, что предложению (1) можно дать иное толкование так, чтобы термин «кошка» в *действительном* мире указывал на *вишни*, а «коврик» — на *деревья* без воздействия на истинностное значение предложения (1) в любом возможном мире. («Находится на» будет сохранять свою первоначальную интерпретацию.)

Идея состоит в том, что предложение (1) получит новую интерпретацию, в которой оно будет означать:

(а) Кошка* на коврикe*.

Определение свойства быть *кошкой** (и, соответственно, быть коврикoм*) дается перечнем случаев его реализации. Эти три случая таковы:

(а) Некоторая кошка находится на некотором коврикe, и некоторая вишня находится на некотором дереве.

(b) Некоторая кошка находится на некотором коврикe, и никакая вишня не находится ни на каком дереве.

(с) Ни один из предшествующих случаев.

Вот определение этих двух свойств:

Определение «кошки»*

x — кошка*, если и только если имеет место (а) и x — вишня; или имеет место (b) и x — кошка; или имеет место (с) и x — вишня.

Определение «коврика»*

x — коврик*, если и только если имеет место (а) и x — дерево; или имеет место (b) и x — коврик; или имеет место (с) и x — кварк.

В возможных мирах, подпадающих под случай (а), предложение «кошка на коврикe» истинно, и «кошка* на коврикe*» также истинно (потому что вишня находится на дереве, а в мирах этого вида все вишни — кошки* и все деревья — коврики*). Поскольку в действительном мире некоторая вишня находится на некотором дереве, действительный мир является именно таким миром, и в действительном мире «кошка*» указывает на вишни, а «коврик*» указывает на деревья.

В возможных мирах, подпадающих под случай (b), «кошка на коврике» истинно, и «кошка* на коврике*» также истинно (потому что в мирах, подпадающих под случай (b), термины «кошка» и «кошка*» коэкстенсивны, как и «коврик» и «коврик*»). (Обратите внимание, что хотя кошки — это кошки* в мирах, подпадающих под случай (b), они *не* являются кошками* в *действительном* мире.)

В возможных мирах, подпадающих под случай (c), «кошка на коврике» ложно, и «кошка* на коврике*» также ложно (потому что вишня не может быть на *кварке*).

Подводя итоги, мы видим, что *в каждом возможном мире* кошка находится на коврике, если и только если кошка* находится на коврике*. Таким образом, мы даем иное толкование слову «кошка», назначая ему интенционал, который мы назначили слову «кошка*» и одновременно даем иное толкование слову «коврик», назначая ему интенционал, который мы назначили слову «коврик*». Результатом этой реинтерпретации будет то, что «кошка на коврике» будет означать то, что по определению предназначено было означать «кошка* на коврике*»; и это было бы совершенно совместимо с тем способом, которым назначаются истинностные значения для «кошка на коврике» в каждом возможном мире.

В приложении я показываю, что более сложная реинтерпретация такого вида может быть выполнена для всех предложений языка в целом. Из этого следует, что всегда есть бесконечно много различных интерпретаций предикатов языка, которые назначают «правильные» истинностные значения предложениям во всех возможных мирах, *независимо от того, как эти «правильные» истинностные значения выделены*. Куайн приводит доводы в пользу подобного заключения в «Слове и объекте»; в примере Куайна (в применении к английскому языку) выражение «это кролик» интерпретируется так, чтобы означать «это частица кролика» (где «ломтик кролика» — трехмерное пространственное сечение целого четырехмерного кролика, локализованного в пространстве-времени), или — еще одна альтернатива — так, чтобы означать «это еще одно воплощение кроличести». (Эта последняя реинтерпретация также дает иное толкование синтаксической форме предложения или, по крайней мере, его логической грамматике.) Куайн делает замечание,

которое я только что сделал и которое состоит в том, что условия истинности для целого предложения оставляют референцию недоопределенной. Поскольку как «частицы кролика», так и «кроликовость», и «неотделенные части кролика» тесно связаны с кроликами, можно вынести из «Слова и объекта» впечатление, что все реинтерпретации, которые оставляют истинностное значение предложения неизменным, по крайней мере близко связаны со стандартной интерпретацией (тем способом, которым части кролика и кроликовость связаны с кроликами). Этот аргумент, разъясненный в приложении и иллюстрированный в этой главе, показывает, что условия истинности для «кошка на коврике» не исключают даже той возможности, что «кошка» указывает на *вишни*.

«ВНУТРЕННИЙ» И «ВНЕШНИЙ»

Возможно, первое, что приходит на ум, когда мы сталкиваемся с ненормативными интерпретациями — вроде той, в которой «кошка» интерпретируется как *кошка**, а «коврик» как *коврик** — это отклонить их как представляющие нам незначительный парадокс. Но *подлинные* парадоксы никогда не бывают незначительными; они всегда показывают, что со способом нашего размышления что-то не в порядке. Возможно, второй реакцией было бы возразить, что *кошка** и *коврик** — «странные» свойства; конечно, наши термины соответствуют «ощутимым» свойствам (вроде *быть кошкой* или *быть ковриком*), а не таким «несерьезным» свойствам, как они. Можно было бы объяснить тот способ, которым *кошка** (или, скорее, *кошачесть**) является несерьезным свойством, отметив, что можно «построить машину», чтобы «осматривать» вещи и «сообщать», действительно ли они — кошки (человек и есть такая «машина»), но нельзя строить машину, которая различит (в любом мире, который напоминает наш по законам и общим условиям), действительно ли это — *кошка**. Если машина (или человек) смотрит на что-то и видит, что это не является ни кошкой, ни вишней, то они могут различить, что это — *не кошка**, но если вещь является кошкой или вишней, то машина или человек должны быть осведомлены об истинностных значениях предложений «кошка на коврике» и «вишня на дереве», чтобы решить, видит ли машина или

человек *кошку**, и эти истинностные значения выходят за пределы того, чему можно научиться, лишь исследуя объект, представленный для осмотра.

К сожалению, можно дать иное толкование «видит» (скажем, *видит**) — такое, что два предложения:

(3) Джон (или кто бы то ни было) видит кошку

и

(4) Джон *видит** кошку*

— будут иметь одинаковое истинностное значение в каждом возможном мире (с помощью метода, приведенного в приложении). Так, всякий раз, когда человек видит кошку, он *видит** кошку*; опыт, который мы обычно имеем, когда видим кошку — это тот же опыт, который мы обычно имеем, когда *видим** кошку*, и так далее. Точно так же мы можем дать иное толкование «осматривает» и «распознает» — такое, что, когда машина осматривает кошку, она *осматривает** кошку*, и когда она «распознает», что это кошка, она *распознает**, что это кошка*.

Используем следующую иллюстрацию (предложенную Нози-ком): предположим, что половина из нас (возможно, женщины) использует слово «кошка» со значением «кошка*», «коврик» со значением «коврик*», «вид» со значением «вид*», «сообщает» со значением «сообщает*», и так далее. Предположим, что другая половина (мужчины) использует слово «кошка», чтобы обозначить кошек, «коврик», чтобы обозначить коврики, «вид», чтобы обозначить вид, и так далее. Как могли бы мы узнать, о чем речь? (Если вы спросите мужчину, на что указывает «кошка», он ответит «на кошек, конечно», и так же ответит женщина, *на что бы ни указывала «кошка»*.)

Дело в том, что сам тот факт, что можно *построить машину, чтобы осматривать вещи и узнавать, если они — кошки*, отличает кошек от кошек*, если можно быть уверенным в том, что «осматривать» и «узнавать» указывают на осмотр и узнавание, а ска-

7. Женщина могла бы ответить, что предположение, что она указывает на кошек*, когда говорит «кошка», *противоречиво* (так как что бы она ни именовала «кошкой» в пределах своего языка, это и является кошкой). Такой ответ не очень облегчает дело; это не исключает возможность того, что то, что она называет кошкой, — это то, что мужчины называют кошкой*, и наоборот; в этом и состоит замечание Нозика.

зять, как устанавливается референция *этих* слов, не легче, чем сказать, как устанавливается референция слова «кошка». Можно сказать, что, когда я смотрю на нечто и думаю, что это — кошка, то мои «ментальные представления», визуальные или осязаемые образы, вербализованная мысль «кошка» и так далее, *указывают* на кошачьсть и на различные другие физические или биологические свойства (иметь определенную форму, определенный цвет, принадлежать к некоторому виду), а не на их аналоги; это может быть истинно, но это лишь повторяет то, что референция устанавливается скорее тем, а не иным путем. Это — то, что мы хотим объяснить, а не разыскиваемое объяснение.

Однако можно было бы возразить, что «*определения «кошки»*» и «*коврика»*», данные выше, указывают на другие вещи, а не на наш предмет (вишни на деревьях и кошки на ковриках), и таким образом означают *внешние* свойства объектов, которые имеют эти свойства. В действительном мире каждая вишня — кошка*; но она не была бы кошкой*, даже при том, что ее внутренние свойства были бы точно теми же, если бы ни на каком дереве не было никакой вишни. И напротив, действительно ли нечто является кошкой — это зависит только от ее внутренних свойств. Позволит ли нам упомянутое здесь различие — различие между *внутренними* и *внешними* свойствами — охарактеризовать и исключить «странные» интерпретации?

Такое решение оказывается неудовлетворительным в силу некоторой симметрии в отношениях «кошки» и «коврика» к «кошке*» и «коврику*». Так, предположим, что мы определяем «вишня*» и «дерево*» таким образом, что в возможных мирах, подпадающих под случай (а), вишни* — это кошки, а деревья* — коврики; в возможных мирах, подпадающих под случай (b), вишни* — это вишни, а деревья* — деревья; и в возможных мирах, подпадающих под случай (с) вишни* — это кошки, а деревья* — фотоны. Тогда мы можем определять «кошку» и «коврик» посредством *-терминов следующим образом. Случаи:

- (а*) Некоторая кошка* находится на некотором коврике*, и некоторая вишня* находится на некотором дереве*.
- (b*) Некоторая кошка* находится на некотором коврике*, и никакая вишня* не находится ни на каком дереве*.
- (с*) Ни один из предшествующих случаев.

Как ни странно, эти случаи — просто наши прежние случаи (а), (б), (с) согласно новому описанию. Теперь мы можем дать определение:

Определение «кошки»

x — кошка, если имеет место (a^*) и x — вишня*; или имеет место (b^*) и x — кошка*; или имеет место (c^*) и x — вишня *. (Обратите внимание, что во всех трех случаях кошки оказываются кошками.)

Определение «коврика»

x — коврик, если и только если имеет место (a^*) и x — дерево *; или имеет место (b^*) и x — коврик*; или имеет место (c^*) и x — кварк*. (Предположив, что кварк* определен так, чтобы в случаях типа (c^*) кварки* являются ковриками, мы получим во всех трех случаях, что коврики оказываются ковриками.)

Заключение, к которому мы приходим, таково: в перспективе языка, принимающего «кошка*», «коврик*» и т. д. за примитивные свойства, «кошка» и «коврик» указывают на «внешние» свойства — свойства, в определениях которых упоминаются другие объекты, нежели x в то время как для «нормального» языка, языка, где «кошка» и «коврик» указывают на кошачесть и ковриковость (Вы знаете, какие свойства я подразумеваю, дорогой читатель!), наоборот, «кошка*» и «коврик*» указывают на «внешние» свойства. Считать что-либо «внутренним» или «внешним» — это зависит от того, какого рода свойства принимаются за *основные*; никакое свойство не является внутренним или внешним само по себе.

«ВЫЖИВАНИЕ» И РАЗВИТИЕ

В настоящее время популярна идея о том, что сам эволюционный процесс так или иначе породил соответствие между нашими словами и ментальными представлениями и внешними вещами; говорят, что мы не *выжили* бы, если бы не было такого соответствия), и что это соответствие и есть, по крайней мере примитивное, отношение референции.

Но какое отношение «соответствие» и «референция» имеют к *выживанию*? В этом плане, какое отношение *истина* имеет к выживанию?

Здесь мнения расходятся. Некоторые философы полагают, что мы не выжили бы, если (достаточно многие из) наших убеждений не были бы *истинны*. Другие философы утверждают, что даже наши лучше всего установленные научные убеждения не истинны, или по крайней мере, что у нас нет никакой причины думать, что они таковы. Томас Кун предположил, что наши убеждения лишь «указывают» на объекты *в пределах* этих убеждений (примерно таким способом, каким «Гамлет» лишь указывает на человека в пьесе); по мнению Куна, успех науки объясняется методом проб и ошибок, а не соответствием между ее объектами и реальными вещами. Бас ван Фраассен в новой книге доказывает, что успешная теория не должна быть истинной, но лишь «наблюдательно адекватной», то есть правильно предсказывающей наблюдение. Он также объясняет успех (или «наблюдательную адекватность») науки как результат метода проб и ошибок.

Если эти философы правы, то сама идея использования эволюции для обоснования убеждения в *объективном* отношении референции оказывается подорванной. Эволюция, согласно таким инструменталистским представлениям, лишь устанавливает соответствие между некоторыми терминами (терминами наблюдения) и «постоянными возможностями ощущения». Такое соответствие — не *референция*, если мы не отказываемся от идеи, что внешние (наблюдаемые) вещи — нечто большее, нежели конструкции из ощущений.

Я полагаю, однако, что правы другие философы — те, кто говорят, что мы не выжили бы, если бы достаточно многие из наших убеждений не были *истинны*.

Причина, по которой я так считаю, состоит в том, что метод проб и ошибок не объясняет, *почему* наши теории «наблюдательно адекватны»; это можно объяснить, только обращаясь к характеристикам *взаимодействия* между человеком и окружающей его средой, которые объясняют, почему *метод проб и ошибок* успешен. (Метод проб и ошибок успешен не во *всех* предприятиях, в конце концов!) Утверждать, что взаимодействие производит в наших умах *ложные* теории, которые лишь *иногда* дают результатом успешные предсказания, значит постулировать полностью необъяснимый ряд *совпадений*. Но как тот факт, что наши убеждения (приблизительно) истинны, объясняет наше выживание?

Некоторые из наших убеждений глубоко связаны с *действием*. Если я верю предложению «Если нажму на эту кнопку, получу что-то очень ценное» (предположив, что я понимаю это предложение обычным образом, или по крайней мере связываю с ним обычное «заключенное в скобки» или «отвлеченное» убеждение), то я протяну руку и нажму кнопку. Назовем убеждения формы «если я выполню действие x , я получу...», где многоточие описывает *цель*, которую имеет действующее лицо, его *директивные убеждения*. Если слишком многие из наших директивных убеждений ложны, мы будем выполнять слишком много неудачных действий; поэтому истинность (достаточно многих из) наших *директивных убеждений* необходима для выживания.

Но и сами наши директивные убеждения выведены из многих других убеждений: убеждений о характеристиках и каузальной силе внешних вещей и убеждений о наших собственных характеристиках и полномочиях. Будь эти убеждения главным образом ложными, разве это не было бы простым совпадением, если бы они тем не менее вели к истинному предсказанию опыта и к истинным директивным убеждениям? Так, поскольку (достаточно многие) наши *директивные убеждения* истинны, и *лучшее объяснение* этого факта состоит в том, что многие из наших других убеждений (составляющие нашу «теорию обыденного мира») являются по крайней мере приблизительно истинным, мы располагаем обоснованием для убеждения в том, что наша теория обыденного мира является по крайней мере приблизительно истинной и что мы не выжили бы, если бы это не имело место.

Вообразим теперь, что некоторые из нас действительно указывают на вещи, которые назначены нашим терминам ненормативной интерпретацией J (описанной в приложении). Эта интерпретация согласуется со стандартной интерпретацией терминов, указывающих на наш понятийный мир, наши ощущения, нашу волю и т. д. Таким образом, предложение «Как мне кажется, я жму на кнопку», понятое в «заключенном в скобки» смысле (как означающее, что я имею некоторый субъективный опыт сознательного нажатия кнопки), имеет не только те же условия истинности, но и ту же *интерпретацию* при J и при «нормальной» интерпретации I , и то же происходит с предложениями типа «Как мне кажется, я получил ожидаемое удовлетворение».

Далее, если достаточно многие из наших директивных убеждений истинны при ненормативной интерпретации *J*, то наши действия будут, конечно, успешны, и мы, конечно, *выживем* (так как если нас не будет в живых, мы не достигнем этих целей) и будем иметь потомство (так как если у нас не будет потомков, *они* не достигнут этих целей). Короче говоря, *J*-истинность (достаточно многих) директивных убеждений столь же хороша для «эволюционного успеха», как и *I*-истинность. Фактически это *I*-истинность, поскольку условия истинности *каждого* предложения (не только директивных убеждений) одинаковы при *I* и при *J*. Мои директивные убеждения не *только* связаны с тем же самым субъективным *опытом* при интерпретации *I* и при интерпретации *J*; они имеют одинаковые условия истинности. С точки зрения (или отсутствия точки зрения) «эволюции», все, что необходимо — это чтобы достаточно многие мои убеждения были истинными при *любой* интерпретации, которая связывает эти убеждения с уместными *действиями*. Эволюция может создать во мне тенденцию иметь *истинные* убеждения (некоторых видов); но это означает лишь, что развитие затрагивает лингвистически или концептуально опосредованное выживание *через* его тенденцию создать в нас системы репрезентаций, чьи предложения или аналоги предложений имеют некоторые условия истинности (и некоторые условия *действия* или «правила языкового выхода»). Но условия истинности *целых* предложений, как было только что показано, не определяют референцию частей предложения (не помогает здесь и добавление «правил языкового выхода», поскольку они сохраняются при *J*). Из этого следует, что просто ошибочно думать, будто эволюция определяет *уникальное* соответствие (или даже разумно узкий диапазон соответствий) между референциальными выражениями и внешними объектами.

ЧИСТЫЕ И НЕЧИСТЫЕ НАМЕРЕНИЯ

Мы видели, что природа не выделяет никакого соответствия между нашими терминами и внешними вещами. Природа заставляет нас обращаться со словами и мысленными знаками таким образом, чтобы достаточно многие из наших директивных убеждений были истинны, и таким образом, чтобы достаточно многие из наших действий вносили вклад в нашу «общую генетичес-

кую пригодность»; но это не определяет референцию. У. В. Куайн настаивал на том, что референция фактически именно и не определена! По его мнению, было бы иллюзией считать, что у терминов нашего языка есть четко определенные предметы. Вот как он говорит об этом.

Рассмотрим снова нашу стандартную регламентированную систему обозначений, со словарем интерпретированных предикатов и некоторым установленным диапазоном значений для переменных квантификации. Предложения этого языка, которые являются истинными, остаются истинными при бесчисленных реинтерпретациях предикатов и пересмотрах диапазона значений переменных. Действительно, для подходящей реинтерпретации предикатов можно использовать любой диапазон той же величины. Если диапазон значений бесконечен, то для этого можно использовать любой бесконечный диапазон; это — теорема Сколема-Левенгейма. Истинные предложения остаются истинными при всех таких изменениях.

Возможно, тогда наш основной интерес относится скорее к истине предложений и к условиям их истинности, чем к референции терминов.

В следующей главе я исследую предложенную здесь альтернативу — отказаться от идеи, которая пока была посылкой всего обсуждения: что слова находятся в некотором взаимно однозначном отношении к (независимым от речи) вещам и множествам вещей. Может показаться, однако, что есть намного более простой выход: почему бы не сказать просто, что референцию наших терминов устанавливают наши неявные или явные *намерения*?

В начале обсуждения в предыдущей главе я отклонил этот вариант как не дающий содержательного ответа на том основании, что наличие намерений (уместного вида) предполагает способность указания. Здесь может быть полезно расширить это краткое замечание.

Проблема состоит в том, что понятия «намерение» и «ментальное состояние» несколько двусмысленны. Назовем ментальное состояние *чистым* ментальным состоянием, если его присутствие или отсутствие зависят только от того, что происходит «внутри» говорящего. Так, действительно ли я испытываю боль, зависит только от того, что происходит «внутри» меня, но дей-

ствительно ли я знаю, что снег бел, зависит не только от того, действительно ли происходит что-либо «внутри» меня (от убеждений или уверенности в том, что снег является белым), но также и от того, действительно ли снег *является* белым, и, таким образом, от чего-то «вне» моего тела и сознания. Итак, *боль* — чистое ментальное состояние, а *знание* — *нечистое* ментальное состояние. В знании есть компонент (чистого) ментального состояния, но есть также компонент, который не ментален ни в каком смысле: это — компонент, который соответствует условию, что то, в чем человек убежден, еще не является знанием, если убеждение не *истинно*. Я не нахожусь в «состоянии» знания, что снег является белым, если я не нахожусь в подходящем чистом ментальном состоянии; но находиться в подходящем чистом ментальном состоянии никогда не *достаточно* для знания, что снег является белым; в этом должен также участвовать мир.

А что с убеждением? Мы определили *заключенное в скобки* убеждение («понятийный мир») так, что обладание заключенным в скобки убеждением в том, что [на столе есть вода] или обладание понятийным миром, в который входит наличие воды на столе — чистое ментальное состояние. Но, исходя из того, что было сказано прежде, *считать, что на столе есть вода* (без всякого «заклучения в скобки») предполагает, что слово «вода» *действительно указывает на воду*, и это зависит от фактического характера некоторых «парадигм», прямых или косвенных причинных отношений с этими парадигмами и так далее. Когда я убежден, что на столе есть вода, мой двойник на Земле-Двойнике имеет то же самое заключенное в скобки убеждение, но не то же самое убеждение, потому что его слово «вода» указывает на смесь воды и спирта, а не на воду. Короче говоря, полагать, что на столе есть вода — *нечистое* ментальное состояние. (Мозги в Бочке не могли бы быть в этом состоянии, хотя могли бы быть в соответствующем «заклученном в скобки» состоянии.)

Что годится для убеждений, годится также и для намерений. Чистые ментальные состояния намерения — например, намерение, чтобы термин «вода» указывал на воду *в чем-либо понятийном мире* — вообще не устанавливают референцию в реальном мире. Нечистые ментальные состояния намерения — например, намерение, чтобы термин «вода» указывал на настоящую воду — *заранее предполагают* способность указать на (реальную) воду.

Некоторые философы предлагают, что убеждение может быть определено в терминах состояния, которое я назвал «заклученным в скобки убеждением», и референции, таким образом:

Джон убежден, что снег бел = Джон убежден, что [снег бел]
(т. е. снег бел в понятийном мире Джона)

и слова «снег» и «белый» в мышлении Джона (или какие бы слова он ни использовал, чтобы выразить это убеждение), указывают на снег и на свойство быть белым, соответственно.

Не принимая это за правильный и полный анализ того, что такое полагать, что снег бел, мы можем принять эту теорию как делающую безусловно верное замечание: *что наличие убеждения заранее предполагает способность осуществлять референцию.* И точно так же намерение предполагает способность осуществлять референцию! Намерения — не ментальные события, которые *являются причиной* того, что слова имеют референцию: референция входит в намерения (в обычном «нечистом» смысле) как их неотъемлемый компонент. Попытка объяснить референцию в терминах (нечистого) намерения содержала бы круг. А проблема того, как *чистые* ментальные состояния намерения, убеждения и т. д. могут (в надлежащем причинном урегулировании) составлять или вызывать референцию — именно та проблема, которую мы нашли настолько загадочной.

ПРОИСХОЖДЕНИЕ ЗАГАДКИ

На первый взгляд, ничто не кажется более очевидным, чем то, что наши слова и ментальные представления *имеют референцию*. Когда я думаю или говорю «кошка только что вышла», эта мысль обычно *о* нашей кошке Митти; слово «кошка» в предложении, которое я мысленно или вслух произношу, *указывает* на множество объектов, членом которого является Митти. Все же мы только что видели, что характер этого отношения «быть о чем-то», или референции, загадочен

Различие между реальным миром и понятийным миром (и коррелятивное различие между убеждениями и заключенными в скобки убеждениями или намерениями и заключенными в скобки намерениями) непосредственно объясняет часть загад-

ки. Причина, по которой для нас так удивительно и тревожно обнаружить, что существуют непреднамеренные «допустимые интерпретации» нашего языка (где под допустимой интерпретацией я имею в виду просто интерпретацию, которая удовлетворяет соответствующим операциональным и теоретическим ограничениям), частично состоит в том, что в «понятийном мире» говорящего не возникает никакой такой «неопределенности». В моем понятийном мире кошки и кошки* весьма отличны друг от друга (на самом деле, в моем понятийном мире кошки* — это *вишни*). «Кошка на коврике» и «кошка* на коврике*» могут быть логически эквивалентны, но они содержат термины с весьма различными понятийными референтами; Таким образом, действительно кажется странным, что можно перепутать в реальном мире референты одного убеждения и референты другого.

Но если число кошек случится равным числу вишен, то из теорем теории моделей следует (как отмечает Куайн в цитированном фрагменте), что возможна такая реинтерпретация полного языка, которая оставляет истинностные значения всех предложений неизменными при перестановке экстенсионалов «кошки» и «вишни». При помощи только что упомянутых методов такая реинтерпретация может быть построена так, чтобы сохранить все операциональные и теоретические ограничения (а при помощи методов, которые мы иллюстрировали на примере «кошка/кошка*», они могут быть расширены так, чтобы давать «интенсионалы», или функции, которые определяют экстенсионал *в каждом возможном мире*, а не только экстенсионалы в действительном мире). Это не противоречит только что сделанным утверждениям о нашем «понятийном мире», или субъективной системе убеждений, по следующей причине: тот факт, что в нашей системе убеждений или «понятийном мире» никакая кошка не является вишней, означает, что в каждой допустимой интерпретации этой системы убеждений (при каждом назначении референтов внешнего мира терминам, образам и другим репрезентациям, которые мы используем в мысли) референты «кошки» и референты «вишни» должны быть непересекающимися множествами. Но непересечение этих множеств сравнимо со следующим (примечательным) фактом: то, что является множеством «кошек» в *одной* допустимой интерпретации, может быть

множеством «вишен» в *другой* (но одинаково допустимой) интерпретации. Если число допустимых интерпретаций — ровно одна, то из того факта, что понятийные кошки полностью отличаются от понятийных вишен, следует лишь то, что понятийные кошки полностью отличаются от реальных вишен. Если же есть больше чем одна допустимая интерпретация целого языка (как это будет иметь место, если допустимые интерпретации выделяются только операциональными и теоретическими ограничениями), то два термина, которые указывают на непересекающиеся множества в *каждой* допустимой интерпретации, могут иметь те же самые потенциальные референты, когда рассматривается вся совокупность *всех* допустимых интерпретаций. Из того факта, что понятийные кошки так отличаются, как только возможно, от понятийных вишен, не следует, что есть определенные непересекающиеся множества кошек-в-себе и вишен-в-себе.

Ситуация осложняется тем, что операциональные и теоретические ограничения являются естественным способом, чтобы позволить действительному эмпирическому контексту определять допустимую интерпретацию (или интерпретации) системы репрезентаций. Такие ограничения могут до некоторой степени определять, какие предложения языка истинны и какие ложны, однако при этом сохраняется инертность между условиями истинности и референцией.

Куайн, как мы отметили, настроен смириться с этой инертностью и просто подтвердить, что референция *не* определена. Молодой философ Хартри Филд⁸ недавно предложил другой подход. Согласно Филду, референция — «физикалистское отношение», то есть сложное причинное отношение между словами или ментальными представлениями и объектами или множествами объектов. Как предлагает Филд, обнаружение этого физикалистского отношения — задача именно эмпирической науки.

Однако с этим предложением также возникает проблема. Предположим, что возможно натуралистическое или физикалистское *определение* референции, как утверждает Филд. Предположим, что истинно

8. *Field H. Tarski's Theory of Truth // The Journal of Philosophy. Vol. 69.* Взгляды Филда обсуждаются в моей книге: *Putnam H. Meaning and the Moral Sciences.* Routledge and Kegan Paul, 1978.

(1) *x указывает на y*, если и только если *x находится в отношении R к y*,

где *R* — отношение, определяемое в словаре естествознания без того, чтобы использовать любые семантические понятия (то есть без того, чтобы использовать «указывает на» или любые другие слова, которые немедленно делали бы определение круговым). Если (1) истинно и поддается проверке опытным путем, то (1) — предложение, которое является самостоятельно истинным даже в той теории, согласно которой референция устанавливается настолько (и *лишь* настолько), насколько она определяется операциональными и теоретическими ограничениями. (1) — предложение, которое было бы частью нашей теории мира «рефлексивного равновесия» или «идеального предела».

Однако если референция определена лишь операциональными и теоретическими ограничениями, то референция «*x находится в отношении R к y*», *самостоятельно* не определена, и поэтому знание того, что (1) истинно, нам не поможет. Каждая допустимая модель нашего объектного языка будет соответствовать модели нашего метаязыка, в котором справедливо (1); интерпретация «*x находится в отношении R к y*» установит интерпретацию «*x указывает на y*». Но это будет лишь отношением *в каждой допустимой модели*; это не поможет сократить число допустимых моделей вообще.

Это, конечно, вовсе не то, что Филд имеет в виду. Вот что требует Филд: (*a*) есть определенное уникальное отношение между словами и вещами или множествами вещей; и (*b*) именно это отношение нужно использовать как отношение референции в назначении истинностного значения самому (1). Но, как мы только что видели, мы не обязательно выражаем это, лишь *говоря* (1); и остается загадкой, как мы могли бы *научиться выражать* то, что хочет сказать Филд.

Отложив в сторону эту последнюю загадку, рассмотрим взгляд, согласно которому (1), понятое в соответствии с желаниями Филда (т. е. как описывающее определенное уникальное отношение между словами и их референтами), истинно. Если так понятое (1) истинно, то что *делает* его истинным? Учитывая, что между словами и вещами есть много «соответствий», и много даже таких, которые удовлетворяют нашим ограничени-

ям, что *выбирает* одно специфическое соответствие R ? Это не эмпирическая правильность (1), так как это — вопрос наших операциональных и теоретических ограничений. И это, как мы видели, не наши намерения (скорее R участвует в определении того, что означают наши намерения). Кажется, что факт, что R является референцией, должен быть *метафизически необъяснимым* фактом, своего рода первичной, иррациональной, метафизической истиной.

Этот вид первичной, иррациональной, метафизической истины, если таковой возможен, не следует путать с «метафизически необходимой» истиной, недавно представленной Солом Крипке⁹.

Замечание Крипке тесно связано с замечаниями, сделанными выше о референции терминов естественных видов (например, терминов для видов животных, растений и минералов,). Оно состоит в следующем. Если *дано*, что, фактически,

(2) Вода — H_2O

(то есть дано, что (2) истинно в действительном мире), и дано, что (как указывает Крипке) говорящие *подразумевают*, что термин «вода» указывает на только те вещи, которые имеют то же закономерное поведение и тот же конечный состав, что и различные стандартные образцы настоящей воды (то есть говорящие имеют такие намерения даже при разговоре о гипотетических случаях или «возможных мирах»), то из этого следует, что (2) должно также быть истинно в каждом возможном мире, так как описывать гипотетическую жидкость, которая не является H_2O , но которая в некоторых отношениях подобна воде — это значит лишь описывать гипотетическую жидкость, которая *напоминает* воду, а не описывать возможный мир, в котором *вода* — не H_2O . «Метафизически необходимо» (истинно во всех возможных мирах), что вода является H_2O ; но эта «метафизическая необходимость» объясняется вполне земной химией и обыденными фактами о намерениях говорящих осуществить референцию.

Если есть определенное физикалистское отношение R (опре-

9. См. *Kripke S. Naming and Necessity*. Harvard University Press, 1980. (Первоначально изложено на лекциях в 1970 году).

делимо ли оно на языке естествознания в конечном количестве слов или нет), которое и *является* референцией (независимо от того, как мы описываем это отношение, или описываем ли мы его вообще), *этот* факт самостоятельно не может быть следствием наших намерений осуществить референцию; скорее, как мы неоднократно отмечали, он участвует в определении того, что означают сами наши намерения осуществить референцию. Подход Крипке, согласно которому «вода — H_2O » истинно во всех возможных мирах, может быть правильным, даже если референция в действительном мире установлена только операциональными и теоретическими ограничениями; этот подход заранее предполагает понятие референции, а не сообщает нам, является ли референция определенной или что такое референция.

По моему мнению, считать, что некоторое соответствие в силу своих внутренних свойств и *является* референцией (не в результате наших операциональных и теоретических ограничений или наших намерений, но как *окончательный* метафизический факт), значит склоняться к волшебной теории референции. Сама референция при таком подходе становится тем, что Локк называл «субстанциальной формой» (сущностью, которая *внутренне* связана с некоторым именем). Даже если мы хотим рассматривать такие необъяснимые метафизические факты, эпистемологические проблемы, которые сопровождают такой метафизический подход, кажутся непреодолимыми. Поскольку, принимая мир независимых от сознания, независимых от речи объектов (а это — исходная предпосылка того подхода, который мы обсуждаем), мы получаем, как мы видели, много различных «соответствий», которые представляют возможность или являются кандидатами в отношения референции (на самом деле, бесконечно много, если во вселенной бесконечно много вещей). Даже требование, что (1) должно быть истинно, какое бы понятие истины ни соответствовало метафизически выделенному «реальному» отношению референции, не исключает ни одного из этих кандидатов, если само (1) *эмпирически* приемлемо (приемлемо согласно нашим операциональным и теоретическим ограничениям), как мы видели. Но тогда есть бесконечно много *различных* возможных «иррациональных метафизических истин» вида « R — *действительное* (метафизически выделенное) отношение референции». Если сторонник такого подхода допус-

кает вероятность того, что его подход не *вполне* правилен, и что референция может быть метафизически выделена без полного *определения* (метафизически выделенное R может допускать множество приемлемых интерпретаций), то вероятно даже то, в конце концов, что метафизически корректен подход, основанный на операциональных и теоретических ограничениях! Ведь вполне может быть иррациональным метафизическим фактом то, что референция *является* отношением: x указывает на y по меньшей мере в одной приемлемой модели M . Заметим, что все это бесконечное множество метафизических теорий совместимо с истинностью *одних и тех же* предложений, с одной и той же «теорией мира» и той же оптимальной методологией обнаружения истины!

3. ДВЕ ФИЛОСОФСКИЕ ПЕРСПЕКТИВЫ

Проблемы, которые мы обсуждали, обычно приводят к появлению двух философских точек зрения (или, как я назвал их во введении, двух философских темпераментов). Именно этими точками зрения, а также теми следствиями, которые они имеют почти для каждой философской проблемы я и намереваюсь заняться. Проблема «Мозгов в бочке» не представляла бы для нас интереса, — разве что только как логический парадокс, — если бы она не имела прямого отношения к выявлению различия между этими философскими перспективами.

Одной из них является перспектива метафизического реализма. Согласно этой перспективе, мир состоит из некоторой определенно установленной совокупности независимых от сознания объектов. Существует только одно истинное и полное описание «того способа, каким существует мир»¹. Истина представляет собой особого рода отношение соответствия между словами или знаками-мыслями и внешними вещами и совокупностями вещей. Я буду называть эту перспективу *экстерналистской* перспективой, поскольку ее излюбленной точкой зрения является точка зрения Божественного Взора.

Та перспектива, которую я намереваюсь защищать, не имеет какого-либо однозначного названия. Она появилась достаточно поздно в истории философии, и даже в настоящее время ее продолжают путать с другими точками зрения совершенно иного рода. Я буду ссылаться на нее как на *интерналистскую* перспективу, поскольку отличительной чертой этой перспективы является убеждение, что вопрос: «*Из каких объектов состоит мир?*» можно осмысленно задавать только *в рамках* какой-либо теории или описания. Многие философы-«интерналисты», хотя и не все, считают также, что не существует одной-единственной «истинной» теории или описания мира. Согласно интерналистской

1. Скрытая цитата, которая отсылает к названию статьи Нельсона Гудмена «Способ, каким существует мир» («The way the world is»), являющуюся ярким манифестом интерналистского направления современной аналитической философии. — *Прим. пер.*

точке зрения, «истина» представляет собой своего рода разновидность (идеализированной) рациональной приемлемости, т. е. некую разновидность идеальной согласованности (*coherence*) наших убеждений друг с другом и данными нашего опыта, *в той степени, в какой эти данные репрезентированы в нашей системе убеждений*, а отнюдь не соответствие независимым от сознания или речи «положениям дел». Не существует точки зрения Божественного Взора, которую мы могли бы знать или же могли бы с пользой для себя вообразить; имеются исключительно точки зрения действительно существующих личностей, отражающие различные интересы и цели, которым служат их описания и теории. («когерентная теория истины», «анти-реализм», «верификационизм», «плюрализм», «прагматизм», — все это термины, применявшиеся для обозначения интерналистской перспективы; однако каждый из этих терминов отягощен коннотациями, которые являются неприемлемыми в силу своего использования в прошлом).

Философы-интерналисты отвергают гипотезу «мозга в бочке». Для нас гипотеза «мозга, находящегося в мире-бочке» есть только *история*, лингвистическая конструкция, а отнюдь не какой-то возможный мир. Идея, согласно которой эта история могла бы быть истинной в рамках некоторой вселенной, в рамках некоторой параллельной реальности, с самого начала допускает, как нетрудно увидеть, точку зрения Божественного Взора. Ведь *с чьей точки зрения рассказывается эта история?* Ясно, что отнюдь не с точки зрения каких-либо наделенных органами чувств существ, живущих в этом мире. Не ведется это повествование также и с точки зрения какого-либо наблюдателя из другого мира, взаимодействующего с этим миром; поскольку «мир» по определению включает в себя все, что каким-либо образом взаимодействует с вещами, которые этот мир содержит. Если *вы*, например, были единственным наблюдателем, который *не* был мозгом в бочке и который при этом шпионил за мозгами в бочке, то мир не был бы миром, в котором *все* наделенные органами чувств существа были бы мозгами в бочке. Так что предположение, что мог бы существовать мир, в котором *все* наделенные органами чувств существа являются мозгами в бочке, с самого начала предполагает существование точки зрения Божественного Взора на истину или, точнее, точки зрения Отсут-

ствия Взгляда на истину (*No Eye view*), т. е. точки зрения на истину как на величину, совершенно не зависимую от наблюдателя.

С другой стороны, для философа-экстерналиста гипотеза, что все мы являемся мозгами в бочке, не может быть так легко отброшена. Поскольку истинность теории заключается не в ее адекватности миру, каким он представляется какому-то наблюдателю или наблюдателям (истина в этом смысле не является «реляционной»), но в ее соответствии миру, каким он существует сам по себе. Проблема же, которую я ставлю перед философом-экстерналистом, заключается в том, что само отношение соответствия, от которого, согласно его точке зрения, зависят истина и референция, не может быть логически для него доступным, если он *является* мозгом в бочке. Так что, если мы *являемся* мозгами в бочке, то мы не можем думать, что мы ими являемся, разве что только *в заключенном в скобки смысле*: (мы есть мозги в бочке); и эта поставленная в скобки мысль не располагает условиями референции, которые сделали бы ее *истинной*. Так что помимо всего прочего просто невозможно, чтобы мы были мозгами в бочке.

Предположим, что мы допускаем «магическую теорию референции». Например, мы могли бы допустить, что некоторые таинственные лучи — назовем их «ноэтическими лучами»² — связывают слова и знаки-мысли со своими референтами. В таком случае никакой проблемы не существует. Мозг в бочке может мыслить слова «Я — мозг в бочке», и когда он делает это, то слово «бочка» соответствует (при помощи ноэтических лучей) реальным внешним бочкам, а слово «в» соответствует (при помощи ноэтических лучей) реальной пространственной локализации. Однако ясно, что такая точка зрения является несостоятельной. Ни один современный философ не согласился бы с этой точкой зрения. Случай с мозгом в бочке представляет для современного реалиста трудную проблему как раз потому, что он хочет иметь теорию истины как соответствия, *не веря* при этом в «ноэтические лучи» (или не веря в само-идентифицирующиеся объекты³, — т. е. в объекты, которые по самому своему

2. «Ноэтические лучи» были подсказаны мне Земахом.

3. Термин «само-идентифицирующиеся объекты» заимствован из книги Дэвида Уиггинса (*Wiggins D. Substance and Sameness. Blackwell, 1980*).

существо соответствуют скорее одному слову или мысли-знаку, нежели другому).

Как мы видели, проблема заключается в следующем. Там существуют эти объекты, которые находятся вне нас. Здесь существует сознание/мозг, осуществляющее мышление/вычисление. Как символы мыслителя (или символы его сознания/мозга) вступают в единственное в своем роде соответствие с объектами и множествами внешних объектов?

Ответ, популярный в настоящее время среди экстерналистов, заключается в том, что хотя ни один знак и в самом деле не соответствует *необходимым образом* скорее одному множеству вещей, чем другому, тем не менее *контекстуальные* взаимосвязи между знаками и внешними вещами (в особенности, каузальные взаимосвязи) дадут возможность объяснить природу референции. Однако это замечание не работает. Например, главной причиной моей веры в электроны являются, вероятно, различные учебники. Однако в тех случаях, когда я произношу слово «электрон», я не *ссылаюсь* на учебники, хотя употребление мною слова «электрон» в этом смысле тесно связано с учебниками. Объекты, выступающие в роли основной причины того, что я обладаю убеждениями, содержащими определенный знак, могут вовсе не быть референтами этого знака.

На это экстерналист ответит, что слово «электрон» не связано с учебниками каузальной цепочкой *соответствующего типа*. (Но как мы можем располагать интенциями, обуславливающими, какие каузальные цепочки относятся к «соответствующему типу», не будучи уже в состоянии *употреблять* данное выражение *референциально*?).

Для интерналиста вроде меня ситуация является принципиально иной. Интерналистская точка зрения также признает, что знаки не соответствуют объектам каким-либо внутренним образом, независимо от того, как и кем они употребляются. Однако знак, который действительно употребляется определенным образом определенным сообществом лиц, его использующих, может соответствовать определенному объекту *в рамках концептуальной схемы этих лиц*. «Объекты» не существуют независимо от концептуальных схем. Мы разделяем мир на объекты, вводя ту или иную схему описания. Поскольку как объекты, так и знаки

в равной степени *соотносятся* со схемой описания, т. е. возможность указать, что чем отображается.

Конечно же, тривиально вести речь о том, на что указывает слово *в рамках* языка, к которому оно принадлежит, используя само это слово. На что указывает слово «кролик»? Как на что, конечно же, на кроликов! На что указывает слово «внеземное существо»? На внеземные существа (если таковые существуют).

Конечно же, экстерналист согласен с тем, что объемом слова «кролик» является множество кроликов, а объемом слова «внеземное существо» — множество внеземных существ. Однако он отказывается считать, что такого рода высказывания сообщают нам, что *собой представляет* референция. Для него раскрытие того, что *собой представляет* референция, т. е. того, какова *природа* «соответствия» между словами и вещами, является настоящей проблемой. (Насколько она настояща для него, мы видели в предыдущей главе). На мой взгляд, мало что можно сказать о референции в рамках концептуальной системы, кроме упомянутых чуть выше тавтологий. Идея о том, что каузальная связь является необходимой, опровергается тем фактом, что слово «внеземное существо» несомненно указывает на внеземные существа независимо от того, вступали ли мы когда-нибудь в каузальное взаимодействие с каким-либо внеземным существом или нет!

Философ-экстерналист ответил бы на это, однако, что мы можем указывать на внеземные существа, даже если мы никогда (насколько это нам известно) не встречали ни одного внеземного существа, потому что мы сталкивались с *землянами* и нам известны случаи отношения «не с той же планеты, что» и случаи свойства «разумное существо». И мы можем *определить* внеземное существо как разумное существо, которое не с той же планеты, что и земляне. К тому же выражение «не с той же планеты, что» может быть проанализировано в терминах «не из того же самого места, что» и «планета» (которые могут быть подвергнуты дальнейшему анализу). Таким образом, экстерналист отбрасывает требование, чтобы мы имели какую-то «реальную» взаимосвязь (например, каузальную взаимосвязь) со *всем тем*, на что мы можем референциально указывать, и требует только того, чтобы *базисные* термины референциально указывали на те виды вещей (и отношений), с которыми мы состоим в

реальной взаимосвязи. А затем, как он утверждает, используя базисные термины в сложных комбинациях, мы можем строить дескриптивные выражения, которые осуществляют референцию к тем видам вещей, с которыми мы не находимся в реальной взаимосвязи, и которые даже могут не существовать, как, например, внеземные существа.

В действительности, уже относительно таких простых слов, как «лошадь» или «кролик» философ-экстерналист мог бы подметить такую особенность, что их объем включает в себя множество вещей, с которыми мы не находимся в отношении каузального взаимодействия (например, тех лошадей и кроликов, которые будут существовать в *будущем*, или же таких лошадей и кроликов, которые никогда не встречались с каким-нибудь человеческим существом). Когда мы используем слово «лошадь», то осуществляем референцию не только к тем лошадям, с которыми мы находимся в реальной взаимосвязи, но также и ко всем остальным вещам *того же самого вида*.

Однако в данном случае мы должны заметить, что выражение «того же самого вида» не имеет смысла вне рамок категориальной системы, указывающей, какие свойства считаются сходными, а какие — нет. Помимо всего прочего, в *известном* смысле все относится «к тому же самому виду», что и другое. Вся эта запутанная история в целом, рассказывающая, как мы осуществляем референцию к определенным вещам благодаря тому факту, что они связаны с нами «каузальными цепочками соответствующего типа», а также как мы осуществляем референцию к другим вещам, которые относятся к «тому же самому виду», что и вещи, которые связаны с нами каузальными цепочками соответствующего типа, а также к еще и другим вещам «при помощи описания», является не столько ложной, сколько бесполезной. Лошади, с которыми я не встречался, относятся к «тому же самому виду», что и лошади, с которыми я *встречался*, исключительно потому, что как первые, так и вторые — *лошади*. Опять-таки та формулировка проблемы, которая дается метафизическим реалистом, создает ложное впечатление, будто бы мне приходится начинать со всех этих объектов самих по себе, что затем мне как будто бы удастся заарканить какую-то сравнительно небольшую их часть (ее составляют лошади, с которыми я нахожусь в «реальной» взаимосвязи благодаря «каузальной це-

почке соответствующего типа»), а затем я сталкиваюсь с проблемой, как добиться того, чтобы мое слово («лошадь») охватывало не только те объекты, которые я «заарканил», но также и те, которые мне заарканить не удалось, либо потому, что они слишком отдалены от меня в пространстве и во времени, либо по какой-то иной причине. И «решение» этой псевдо-проблемы, как я его себе представляю, — а именно то «решение», которое предлагается метафизическим реалистом, — заключается в утверждении, что слово *автоматически* охватывает не только те объекты, которые я заарканил, но также и объекты, которые относятся к *тому же самому виду* — к тому же виду *сами по себе*. Однако в таком случае по существу утверждается, помимо всего прочего, что мир содержит само-идентифицирующиеся объекты, поскольку именно это и подразумевается утверждением, что *мир*, а не мыслитель, распределяет вещи по отдельным видам.

В известном смысле и я стал бы утверждать, что мир *действительно* состоит из «само-идентифицирующихся объектов», однако это вовсе не тот смысл, который приемлем для экстерналиста. Если, как я считаю, сами объекты являются столь же конструируемыми, сколь и открываемыми, если они являются продуктами нашего концептуального изобретения не в меньшей степени, чем «объективным» фактором опыта, т. е. фактором, независимым от нашей воли, то, конечно же, по самому своему существу объекты группируются под определенными этикетками; ведь эти этикетки представляют собой средства, которые мы используем прежде всего для того, чтобы строить версию мира, состоящего из таких объектов. Однако *этот* вид «само-идентифицирующегося объекта» не является независимым от сознания; экстерналист же хочет мыслить мир состоящим из объектов, которые *в одно и то же время* являются и независимыми от сознания, и само-идентифицирующимися. А это невозможно.

ИНТЕРНАЛИЗМ И РЕЛЯТИВИЗМ

Интернализм не является легковесным релятивизмом, заявляющим, что «годиться все». Отрицать, что имеет смысл задаваться вопросом, «отображают» ли наши понятия что-то, совершенно не затронутое концептуализацией, — это одно; однако счи-

тать. на этом основании, что любая концептуальная система столь же хороша, как и любая другая — это нечто совсем иное. Например, предположим, что какой-то не слишком умный человек воспринял эту идею всерьез и предложил бы такую теорию, которая утверждает, что человек способен летать без помощи технических средств. Если бы он попробовал применить свою теорию на практике, выпрыгнул бы ради этого в окно и чудом остался в живых, то он вряд ли после этого стал бы придерживаться этой теории. Интернализм не отрицает того, что в отношении знания играют роль опытные *исходные данные*, знание не является рассказом, который не имеет иных ограничивающих условий, кроме *внутренней* согласованности; однако он и в самом деле отрицает, что существуют такие исходные данные, *которые сами не формировались бы до известной степени нашими понятиями*, тем словарем, который мы используем для того, чтобы фиксировать и описывать их, или же что существуют исходные данные, *которые допускают одно и только одно описание, независимо от всех концептуальных предпочтений*. Даже наше описание наших собственных ощущений, которое было — в качестве исходной точки знания — столь дорого сердцу целых поколений эпистемологов, испытывает мощное воздействие (как и наши ощущения, коли на то пошло) множества наших концептуальных предпочтений. Сами исходные данные, на которые опирается наше знание, являются концептуально инфицированными; однако лучше иметь инфицированные исходные данные, чем вообще не иметь никаких данных. Если бы инфицированные данные — это всё, чем мы располагаем, даже в этом случае все то, что нам доступно, сохранило бы свою значимость.

Высказывание, или целая система высказываний, — т. е. теория или концептуальная схема, — становятся рационально приемлемыми в значительной степени благодаря своей согласованности и пригодности; благодаря согласованности «теоретических» или менее опытных убеждений друг с другом и с более опытными убеждениями, а также благодаря согласованности опытных убеждений с теоретическими убеждениями. Согласно тому взгляду, который я буду развивать, наши понятия согласованности и приемлемости тесно переплетаются с нашей психологией. Они зависят от нашей биологии и нашей культуры; они никоим образом не являются «свободными от ценностей». Но

они *суть* наши понятия, и притом понятия чего-то реального. Они определяют своеобразную объективность, *объективность для нас*, даже если она не является метафизической объективностью Божественного Взора. Говоря по-человечески, объективность и рациональность, — это то, чем мы располагаем; а это лучше, чем ничего.

Огрицать идею, что существует когерентная «внешняя» перспектива, т. е. теория, которая просто истинна «сама по себе», безотносительно к каким-либо возможным наблюдателям, не означает *отождествлять* истину с рациональной приемлемостью. Истина не может быть отождествлена с рациональной приемлемостью по одной простой причине: истина считается свойством высказывания, и как таковая она не может быть потеряна, тогда как обоснование (*justification*) — может. Высказывание «Земля — плоская» было, что весьма вероятно, рационально приемлемо 3000 лет тому назад; однако оно рационально неприемлемо в настоящее время. Однако было бы ошибкой утверждать, что высказывание «Земля — плоская» было *истинно* 3000 лет тому назад; поскольку это означало бы, что форма Земли изменилась. В действительности рациональная приемлемость и инициируется личностью, и соотносится с ней. Вдобавок к этому, рациональная приемлемость есть дело степени; об истине тоже иногда говорят как о деле степени (например, мы иногда говорим, что выражение «Земля представляет собой шар» *приблизительно истинно*); однако под «степенью» в данном случае имеется в виду *точность* высказывания, а не степень приемлемости или обоснованности.

Вышеприведенные соображения, на мой взгляд, свидетельствуют не о том, что точка зрения экстерналиста все же является истинной, но что истина представляет собой *идеализацию* рациональной приемлемости. Мы рассуждаем так, как если бы идеальные с точки зрения эпистемологии условия и в самом деле имели место и мы называем высказывание «истинным», как если бы оно было обосновано в подобного рода условиях. «Эпистемологически идеальные условия» чем-то напоминают «плоскости, лишенные трения»: в действительности мы не можем достичь эпистемологически идеальных условий, или даже быть абсолютно уверенными в том, что мы достаточно к ним приблизились. Однако в действительности нельзя создать и плоскости,

лишенные трения, и все-таки разговор о плоскостях, лишенных трения, имеет свою «наличную стоимость», поскольку мы можем приблизиться к ним в очень высокой степени.

Вероятно, может создаться впечатление, что объяснение истины в терминах обоснования при идеальных условиях представляет собой объяснение ясного понятия при помощи терминов смутного понятия. Однако «истинно» *не* является столь ясным, как только мы отходим от таких заезженных примеров, как «снег бел». В любом случае, я пытаюсь дать не *формальное* определение истины, но неформальное разъяснение этого понятия.

Если сравнение с плоскостями, лишенными трения, оставить в стороне, то к числу двух ключевых идей теории истины как идеализации относится (1) то, что истина независима от обоснования здесь и сейчас, но не может считаться независимой от *любых* обоснований. Утверждать, что высказывание истинно, означает утверждать, что оно могло бы быть оправдано. (2) Вторая важная идея сводится к тому, что истина считается чем-то устойчивым и «непротиворечивым»; если и высказывание, и его отрицание могли бы быть «оправданы» даже при самых идеальных условиях, то нет никакого смысла утверждать, что такое высказывание *имеет* истинностное значение.

ТЕОРИЯ «ПОДОБИЯ»

Теория, согласно которой истина есть соответствие, является достаточно естественной. Возможно, до Канта вообще нельзя отыскать *какого-либо* философа, который *не* придерживался бы корреспондентской теории истины.

Недавно⁴ Майкл Даммит провел различие между *не-реалистической* (т. е. той, что я называю «интерналистской») и *редукционистской* точками зрения для того, чтобы указать, что редукционисты могут быть мегафизическими реалистами, т. е. приверженцами корреспондентской теории истины. Редукционизм, если рассматривать его с точки зрения отношения к классу ут-

4. Взгляды Даммита излагаются в статьях «Что такое теория значения I, II», опубликованных в сборнике «Истина и другие загадки» (*Dummit M. Truth and Other Enigmas. Harvard, 1980*). В его готовящихся к публикации (возможно) лекциях, посвященных Уильяму Джеймсу и прочитанных в Гарварде в 1976 году, эти его взгляды развиваются более подробно.

верждений (например, утверждений относительно ментальных событий), представляет собой точку зрения, согласно которой факты, находящиеся за пределами этого класса, «делают истинными» утверждения этого класса. Например, согласно одной из разновидностей редукционизма, факты, связанные с поведением, «делают истинными» утверждения относительно ментальных событий⁵. В качестве другого примера можно привести точку зрения епископа Беркли, согласно которой сфера того, что «реально существует», исчерпывается сознаниями и их ощущениями. Эта точка зрения является *редукционистской*, поскольку Беркли считает, что предложения о столах, стульях и иных обычных «материальных объектах» в действительности делают истинными факты, касающиеся ощущений.

Если точка зрения является редукционистской относительно утверждений одного вида, но настаивает при этом на корреспондентской теории истины применительно к предложениям *редуцирующего* класса, то эта точка зрения есть, в своей основе, точка зрения метафизического реализма. Подлинно не-реалистическая точка зрения является не-реалистической во всех отношениях.

Очень часто делают ошибку, когда считают философов-редукционистов не-реалистами, однако Даммит, конечно же, прав; их разногласия с другими философами касаются того, *что в действительности существует*, а не понятия истины. Если мы избежим этой ошибки, то в этом случае заявление, которое я только что сделал, а именно, что невозможно найти такого философа до Канта, который бы не был метафизическим реалистом, по крайней мере в отношении тех утверждений, которые они считали *базисными* или не поддающимися редукции, будет выглядеть намного более убедительным.

Древнейшей формой корреспондентской теории истины, существующей уже приблизительно 2000 лет, является та, что античные и средневековые философы приписывают Аристотелю. Я не уверен, что Аристотель и в самом деле придерживался ее; однако на это указывает его язык. Я буду называть эту теорию

5. Имеется в виду доктрина бихевиоризма в психологии и философии сознания, согласно которой предметом психологического исследования могут быть только акты поведения человека, доступные для внешнего наблюдения. — Прим. пер.

теорией референции как подобия; поскольку она считает, что отношение между репрезентациями в нашем уме и внешними объектами, на которые эти репрезентации указывают, представляет собой буквальное *подобие*.

Эта теория, как и современные теории, использует идею ментальной репрезентации. Это представление, т. е. образ внешней вещи, который есть у ума, Аристотель называет *фантасма*, т. е. образ. Отношение между образом и внешним объектом, благодаря которому образ репрезентирует уму внешний объект, состоит (согласно Аристотелю) в том, что образ *имеет одинаковую* с внешним объектом *форму*. Поскольку образ и внешний объект сходны между собой (имеют одинаковую форму), ум, имея доступ к образу, имеет также и непосредственный доступ к самой *форме* внешнего объекта⁶.

Сам Аристотель говорит, что образ не разделяет с объектом такие свойства, как *краснота* (т. е. краснота в наших умах не является буквально тем же самым свойством, что и краснота объекта), которое может быть воспринято благодаря только одному органу чувств, но разделяет такие свойства, как *длина* или *форма*, которые могут быть восприняты при помощи более чем одного органа чувств (которые являются «общим воспринимаемым» в противоположность «единичным воспринимаемым»).

В XVII веке теория подобия начинает претерпевать ограничения, значительно большие по своим масштабам, чем те, что имели место при Аристотеле. Так, Декарт и Локк считают, что в случае «вторичного качества» такого, как цвет или степень плотности ткани, было бы абсурдно предполагать, что свойство ментального образа является *буквально* тем же самым свойством, что и свойство физической вещи. Локк был сторонником корпускулярной доктрины, т. е. приверженцем атомистической теории материи, и подобно современному физику, он считал, что чувственно данной красноте моего образа красной ткани соответствует не простое свойство ткани, но весьма сложное диспозиционное свойство или «способность»: способность вызывать ощущения именно этой разновидности (ощущения, которые проявляют «субъективно красное», выражаясь языком психофизики). В свою очередь, эта способность имеет свое

6. См.: Аристотель. О душе, книга III, главы 7 и 8.

объяснение, которое не было известно во времена Локка, состоящее в особенной микроструктуре кусочка ткани, благодаря чему он избирательно поглощает и отражает световые волны различной длины. (Этот *вид* объяснения был дан уже Ньютоном). Если мы говорим, что обладание такой микроструктурой означает «бытие красным» в случае с кусочком ткани, то ясно, что какой бы ни была природа субъективно красного, событие в моем уме (или даже в моем мозге), которое происходит тогда, когда я имею ощущение красного, *не* влечет за собой чего-либо «субъективно красного» в моем уме (или мозге). Те свойства физической вещи, которые делают ее частным случаем физически красного, и свойства ментального события, которые делают его частным случаем субъективно красного, совершенно отличны друг от друга. Красный кусочек ткани и красный вторичный образ *не* являются буквально подобными. Они не имеют общей Формы.

Из-за тех свойств (форма, движение, местоположение), которые в силу своей корпускулярной философии Локк был вынужден считать базисными и не поддающимися редукции, он, однако, стремился придерживаться теории референции как подобия. (В действительности, некоторые исследователи Локка в настоящее время спорят по этому поводу; однако Локк и в самом деле утверждал, что в случае первичных качеств имеется «подобие» между идеей и объектом и что «нет подобия» между идеей *красного* или *теплого* и краснотой или теплотой объекта⁷. И то прочтение Локка, которое я описываю, было широко распространено как среди его современников, так и среди читателей XVIII столетия).

TOUR DE FORCE⁸ БЕРКЛИ

Беркли обнаружил одно весьма неприятное следствие теории референции как подобия: она подразумевает, что не существует ничего, кроме ментальных сущностей («духов и их идей», т. е. сознаний и их ощущений). Обычно не принимается во внима-

7. См.: *Locke J. An Essay Concerning Human Understanding. Book II, chapter VIII.*

8. Силовой прием (*фр.*). — *Прим. пер.*

ние то обстоятельство, что посылка, от которой отталкивался Беркли — теория подобия — не была просто заимствована им у Локка (или вложена им в уста Локка), но была общепринятой теорией референции еще задолго до него и потом еще целое столетие после него; однако мы только что отметили, насколько древней была в действительности эта теория.

Аргумент Беркли является достаточно простым. Обычный философский аргумент против теории подобия в случае вторичных качеств является правильным (аргумент от относительности восприятия), но он имеет силу также и в случае первичных качеств. Длина, форма, движение объекта воспринимаются различно различными наблюдателями и одним и тем же наблюдателем в различных обстоятельствах. Спрашивать, имеет ли *стол* ту же самую длину, что и *мой* образ его, или ту же длину, что и *ваш* образ его, — означает задавать абсурдный вопрос. Если длина стола составляет три фута, и я воспринимаю его ясно и отчетливо, то обладаю ли я *ментальным образом в три фута длиной*? Задавать этот вопрос означает отдавать себе отчет в его бессмысленности. Ментальные образы не имеют *физической* длины. Они не могут быть подвергнуты сравнению со стандартным эталоном длины, находящимся в Париже. Физическая длина и субъективная длина должны быть столь же отличными друг от друга, что и физическая краснота и субъективная краснота.

Заключение Беркли можно сформулировать иначе: *Ничто не может быть сходным с ощущением или образом помимо другого ощущения или образа*. Если дано это заключение, а также принято допущение (которое пока еще считается несомненным), что механизм референции представляет собой подобие между нашими «идеями» (т. е. нашими образами) и тем, что они репрезентируют, то отсюда сразу же следует, что ни одна «идея» (ментальный образ) не может репрезентировать что-либо или отсылать к чему-либо помимо другого образа или ощущения. Только феноменальные объекты можно мыслить, постигать, делать предметом референции. И если вы не можете думать о чем-то, то это означает, что вы не можете думать о нем как о существующем. Рассуждение о материальных объектах, если только оно не считается производным от рассуждения о тех регулярностях, что встречаются в наших восприятиях, оказывается совершенно невразумительным.

Тенденция рассматривать Беркли как без пяти минут безумца, упрямящегося в своей неправоте, чье поведение выглядит совершенно возмутительно, пусть даже он при этом и является выдающимся мыслителем, имевшая место как в его время, так и позже, была обязана своим появлением неприемлемости его заключения, что материя в действительности не существует (она существует только как конструкция из ощущений), а отнюдь не неприемлемости каких-то его посылок. Однако тот факт, что можно было вывести столь неприемлемое заключение из теории подобия, вызвал кризис в философии. Философы, которые не пожелали следовать за Беркли по пути субъективного идеализма, должны были придумать референции какое-то другое объяснение.

ИСТОЛКОВАНИЕ ЗНАНИЯ И ИСТИНЫ КАНТОМ

Я собираюсь утверждать, что хотя Кант никогда и не характеризовал свое творчество подобного рода образом, тем не менее, он лучше всего прочитывается как мыслитель, с самого начала предлагающий то, что я назвал «интерналистской» или «внутренне реалистической» точкой зрения на истину.

Начать надо с того, что Кант считал субъективный идеализм Беркли совершенно неприемлемой позицией (о чем он неоднократно говорил), а также считал в равной степени неприемлемым и каузальный реализм — взгляд, согласно которому мы воспринимаем непосредственно только ощущения и *выводим* из них материальные объекты при помощи некоторого весьма проблематичного вида вывода. Взгляд, согласно которому тот факт, что передо мной находится стол, на котором я пишу эти страницы, является не более чем весьма сомнительной гипотезой. Кант называет его «скандалом [в философии]».

Во-вторых, я считаю, что Кант отдавал себе ясный отчет в том, как работает аргумент Беркли: он видел, что он зависит от теории референции как подобия и что опровержение аргумента Беркли требует опровержения этой теории. В данном случае я приписываю Канту взгляд, который он сам никогда не выражал в подобного рода словах (конечно же, речь о «референции» как об отношении между ментальными знаками и тем, что они представляют, является достаточно современной, хотя сама

проблема отношения между ментальными знаками и тем, что они представляют, является достаточно древней). Однако мы увидим, что то, что Кант *в действительности* говорит, приводит к отказу от теории референции как сходства.

Позвольте мне предложить способ прочтения Канта, который может оказаться полезным, хотя он и представляет собой только приближение к правильной интерпретации. Представьте себе, что Кант принимает мнение Беркли, что аргумент от относительности восприятия в той же степени касается так называемых «первичных» качеств, как и вторичных, но что он реагирует на это иначе, нежели Беркли. Припомните, что реакция Беркли состояла в том, чтобы стереть всякое различие между первичными и вторичными качествами и возвратиться к тому, что Локк незадолго до этого назвал «простыми» качествами ощущения. Они и должны были стать теми базисным сущностям, на которые мы можем референциально ссылаться. Припомните, что вывод, к которому пришел Локк в результате рассмотрения вторичных качеств, состоял в том, что мы можем постичь их, — в качестве свойств физического объекта, — только в виде сил, т. е. в качестве свойств — *неспецифицированной природы* — которые заставляют объект воздействовать на нас определенным образом. Утверждать, что нечто является красным, или теплым, или покрытым мехом, означает утверждать, что это нечто является таковым *по отношению к нам*, а не с точки зрения Божественного Взора.

Я предполагаю, что (в первом приближении) способ прочтения Канта заключается в выдвижении догадки, что то, что Локк говорил по поводу вторичных качеств, истинно по отношению ко *всем* качествам — простым, первичным, а также вторичным (конечно, мало проку их различать)⁹.

9. Кант обобщает свою собственную позицию примерно таким же самым образом в своих «Пролегоменах»:

«Еще задолго до времен Локка, но в особенности после него стало общепринятым и признанным мнением, что несмотря на действительное существование внешних вещей, о множестве их предикатов можно сказать: они не принадлежат к этим вещам самим по себе, а только к их явлениям и вне нашего представления не имеют никакого собственного существования. Сюда принадлежат теплота, вкус, цвет и пр.. А что я по важным причинам причислил к простым явлениям сверх этих и остальные качества тел, называемые *первичными*, как то: протяжение, место, пространство вообще со всем к нему принадлежа-

Если *все свойства являются вторичными*, то что из этого следует? Это означает, что *все*, что мы говорим о каком-либо объекте, принимает следующую форму: объект таков, что воздействует на *нас* таким-то образом. И *ничто из того*, что мы говорим по поводу какого-либо объекта, не описывает объект таким, каков он есть «сам по себе», т. е. независимо от его воздействия на *нас*, на существа, имеющие нашу рациональную природу и нашу биологическую конституцию. Из этого следует также, что мы не можем допустить какого-либо сходства («simulitude», если выражаться на том английском, который использует Локк), между нашей идеей объекта и какого угодно рода независимой от сознания реальностью, благодаря которой мы, в конечном счете, располагаем опытом этого объекта. Наши идеи объектов не являются *копиями* независимых от сознания вещей.

Примерно так Кант и описывает ситуацию. Он не сомневается в том, что существует *какая-то* реальность, независимая от сознания; это убеждение по существу своему является для него постулатом разума. Он ссылается на эти независимые от сознания элементы реальности при помощи различных терминов: вещь сама по себе (*Ding an sich*); ноуменальные объекты или *ноумены*; в совокупности, *ноуменальный мир*. Однако по поводу этих ноуменальных вещей мы не в состоянии образовать какое-либо реальное понятие; даже понятие ноуменального мира выступает в роли скорее предела мысли (*Grenz-Begriff*), нежели в роли ясного понятия. В настоящее время понятие ноуменального мира считается своего рода необязательным метафизическим элементом мышления Канта. (Однако возможно, что Кант прав: возможно, мы не в состоянии удержаться от мысли о том, что *каким-то образом* существует независимое от сознания «основание» нашего опыта, даже если попытки вести об этом речь немедленно приводят к бессмыслице).

щим (непроницаемость или материальность, форма и пр.), — так не допускать этого нет ни малейшего основания. И как тот, кто признает, что цвета не суть свойства предмета самого по себе, а только видоизменения зрительного чувства, не может называться за это идеалистом, так и мое учение не может называться идеалистическим за то, что я считаю не одни эти, но *даже все свойства, составляющие созерцаемый образ тела, принадлежащими только к его явлению*» [Ср.: Кант И. Прологомены ко всякой будущей метафизике, могущей возникнуть в смысле науки / Пер. с нем. В. С. Соловьева. М.: Прогресс, 1993. С. 60–61. — *Прим. пер.*].

В то же самое время, разговор об обычных «эмпирических» объектах *не* является разговором о вещах самих по себе, но представляет собой разговор о вещах для нас.

В действительности едва различимая особенность позиции Канта заключается в том, что он считает все эти характеристики в равной степени применимыми как к внешним объектам, так и к ощущениям («объектам внутреннего чувства»). Это может показаться странным: что за проблема, соответствует ли идея ощущению или нет? Однако Кант в данном случае имеет в виду нечто более важное.

Предположим, что я имею ощущение *E*. Предположим, что я описываю *E*; скажем, при помощи утверждения «*E* есть ощущение *красного*». Если «красное» означает всего-навсего *наподобие этого*, то все утверждение в целом означает «*E* есть *наподобие этого*» (говоря о *E*), т. е. *E* подобно *E* — и тогда вообще никакого утверждения не было сделано. Как указывает Витгенштейн, в результате все сводится к бессвязному бормотанию. С другой стороны, если «красное» является подлинным *классификатором*, если я заявляю, что это ощущение *E относится к тому же классу*, что и ощущение, которое я называю «красным» *в другие моменты времени*, то мое суждение выходит за пределы непосредственно данного, за пределы «чистой этости» и включает в себя неявную ссылку на другие ощущения, которыми я не обладаю в данный момент, и на *время* (которое, согласно Канту, не является чем-то ноуменальным, но скорее формой, благодаря которой мы упорядочиваем «вещи для нас»)¹⁰. Вопрос о том, являются ли ощущения, которыми я обладал в разные моменты времени и которые я классифицировал как *ощущения красного*, «реально» (ноуменально) сходными, не имеет смысла; если они кажутся сходными (например, если я *вспоминаю* предыдущие ощущения как сходные с этим налично данным ощущением и *предполагаю*, что будущие ощущения, которые я классифицирую подобным

10. В данном случае я являюсь нарочито анахроничным и описываю позицию Канта при помощи примера, взятого из «Философских исследований» Витгенштейна. Однако пример Витгенштейна имеет глубокие кантовские корни: Гегель, создававший свои работы вскоре после Канта и знавший его учение, в явном виде сформулировал идею, согласно которой каждое суждение, даже суждение по поводу чувственных впечатлений, должно выйти за пределы «данного» для того, чтобы вообще *быть* суждением.

же образом, будут, в свою очередь, выглядеть сходными с этим налично данным ощущением, в том виде, в каком оно будет восминаться мне в будущем), то они сходны для меня.

Кант постоянно и разными словами напоминает, что объекты внутреннего чувства *не* являются трансцендентально реальными (ноуменальными), что они являются «трансцендентально идеальными» (вещами для нас), и что они не являются в большей или меньшей степени непосредственно познаваемыми, чем так называемые «внешние» объекты. Ощущения, которые я называю «красными», можно не в большей степени сравнивать с ноуменальными объектами для того, чтобы увидеть, имеют ли они то же самое ноуменальное свойство, нежели объекты, которые я называю «кусками золота», могут ли непосредственно сравниваться с ноуменальными объектами для того, чтобы увидеть, имеют ли они то же самое ноуменальное свойство.

Основание, в силу которого тезис «все свойства являются вторичными» представляет собой только *первое* приближение к позиции Канта, заключается в следующем: тезис «все свойства являются вторичными» (т. е. *все свойства суть силы*) подразумевает, что утверждать по поводу кресла, что оно сделано из сосны, или из чего-то еще, означает приписывать силу (т. е. способность казаться нам сделанным из сосны) ноуменальному объекту; утверждать по поводу кресла, что оно коричневое, означает приписывать другую силу *тому же самому* ноуменальному объекту; и так далее. Согласно этой точке зрения, был бы один ноуменальный объект, соответствующий каждому объекту в том, что Кант называет «репрезентацией», т. е. один ноуменальный объект, соответствующий каждой вещи для нас. Однако Кант явно *отрицает* это. Именно в этом пункте он практически утверждает, что отбрасывает теорию истины как соответствия.

Однако Кант, конечно же, не утверждает, что он отбрасывает теорию истины как соответствия. Напротив, он утверждает, что истина есть «соответствие суждения своему объекту». Однако это определение Кант называет «номинальным определением истины». На мой взгляд, отождествлять «номинальное определение истины» с тем, что метафизический реалист имеет в виду под «теорией истины как соответствия» было бы серьезной ошибкой. Для того, чтобы ответить на вопрос, придерживался ли Кант того, что метафизический реалист подразумева-

ст под «теорией истины как соответствия», мы должны понять, располагал ли он реалистической концепцией того, что он называл «объектом» эмпирического суждения.

Согласно позиции Канта, любое суждение относительно внешних или внутренних объектов (физических вещей или ментальных сущностей) утверждает, что ноуменальный мир, взятый в качестве целого, таков, что он представляет собой описание, которое построило бы рациональное существо (существо, имеющее нашу рациональную природу), обладающее информацией, доступной для существа, наделенного нашими органами чувств (существо с нашей чувственной природой). В этом смысле суждение приписывает какую-то силу. Однако сила приписывается *всему ноуменальному миру в целом*; вы не должны думать, что в силу того, что в нашем представлении даны кресла, лошади и ощущения, то должны существовать, соответственно, ноуменальные кресла, ноуменальные лошади и ноуменальные ощущения. *Не существует даже взаимнооднозначного соответствия между вещами для нас и вещами самими по себе.* Кант не только отбрасывает любое понятие сходства между нашими идеями и вещами самими по себе; он отбрасывает даже любое понятие абстрактного изоморфизма. А это означает, что в рамках его философии нет места для теории истины как соответствия.

Что же собой представляет, в таком случае, истинное суждение? Кант и в самом деле убежден в том, что мы обладаем *объективным* знанием: мы знаем законы математики, законы геометрии, законы физики и множество высказываний относительно индивидуальных объектов — эмпирических объектов, вещей для нас. Употребление термина «знание» и употребление термина «объективный» сводится к утверждению, что *все еще существует понятие истины*. Но что такое истина, если она не является соответствием тому способу, каким вещи существуют сами по себе?

Как я сказал, единственный ответ, который можно извлечь из сочинений Канта, заключается в следующем: фрагмент знания (т. е. «истинное высказывание») есть высказывание, которое рациональное существо приняло бы на основании достаточного опыта такой разновидности, которая доступна для существ, обладающих нашей природой. «Истина» в любом ином смысле недоступна и непостижима для нас. *Истина есть предельное благо пригодности.*

ЭМПИРИСТСКАЯ АЛЬТЕРНАТИВА

В той степени, в которой наш аргумент оказывается подходящим, у философа все еще сохраняется возможность не отказываться от теории истины как соответствия и от теории референции как подобия, *ограничив область их действия ощущениями и образами*. И даже после Канта многие философы продолжали верить в то, что именно подобие и *есть* тот механизм, благодаря которому мы способны иметь идеи, которые отсылают к нашим собственным ощущениям (и, хотя это и более спорно, к ощущениям других людей), и что с эпистемологической точки зрения это и является первичным случаем референции.

Для того чтобы понять, что это соображение не работает, припомните, что сердцевину аргумента Беркли составляло убеждение, что ничто не может быть сходным с «идеей» (ощущением или образом), кроме другой «идеи»; т. е. что не может существовать сходства между ментальным и физическим. Согласно Беркли, наши идеи могут быть похожими на другие ментальные сущности, однако они не могут быть сходными с «материей».

Здесь надо остановиться и дать себе отчет в том, что в одном очень важном отношении это убеждение является ложным. В действительности, *все является сходным со всем в бесконечно большом множестве отношений*. Например, мое ощущение от печатной машинки в данный момент и от монеты достоинством в 25 центов в моем кармане сходны в том отношении, что некоторые из их свойств (ощущение происходит в данный момент, и монета находится в моем кармане в данный момент) являются *результатами моих прошлых действий*; если бы я не сел печатать на машинке, то я не имел бы ощущения; и монета не находилась бы в моем кармане, если бы я не положил ее туда. Как ощущение, так и монета существуют в двадцатом веке. Как ощущение, так и монета описываются на английском языке. И так далее и тому подобное. Число сходств, которые можно найти между любыми двумя объектами, ограничивается исключительно изобретательностью и временем.

Конечно же, в каком-то отдельном контексте слово «сходство» может иметь более ограничительное значение. Однако просто спрашивать «являются ли *A* и *B* схожими?» в тех случаях, когда мы не определили, явно или неявно, какого *вида* сход-

ство имеется в виду, означает задавать бессодержательный вопрос.

Из этого простого факта следует, что идя, согласно которой сходство представляет собой приватный механизм референции, должна приводить к регрессу в бесконечность. Предположим, следуя примеру, предложенному Витгенштейном, что кто-то пытается изобрести «индивидуальный язык», т. е. язык, который референциально отсылает к его собственным ощущениям в том виде, в каком они непосредственно ему даны. Он концентрирует свое внимание на ощущении X и вводит знак E , который он намеревается применять исключительно к тем сущностям, которые качественно тождественны с X . В результате он имеет в виду то, что знак E следовало бы применять ко всем тем сущностям, которые *сходны* с X , и только к ним.

Если этим и исчерпывается *все* то, что он имел в виду. — если он не указывает тот *аспект*, в отношении которого нечто должно иметь сходство с X для того, чтобы подпадать под классификацию E , — то его интенция пуста, как мы только что видели. Поскольку *все что угодно* сходно с X в *каком-то* отношении.

Если же, с другой стороны, он *указывает* на определенный аспект: если он думает, что *ощущение есть E тогда и только тогда, когда оно сходно с X в отношении R* , тогда, поскольку он способен думать подобным образом, он *в то же самое время* способен референциально ссылаться на ощущения, для которых он пытается ввести термин E , и на соответствующее свойство этих ощущений! Но как он оказался в состоянии сделать *это*? (Если мы ответим: «Сосредоточив свое внимание на двух других ощущениях, Z , W , и думая, что эти два ощущения сходны в отношении R тогда и только тогда, когда они сходны с Z , W », то вовлечемся в регресс в бесконечность»).

Затруднение, связанное с теорией референции как подобия — того же рода, что и затруднение, связанное с теорией «каузальной цепи соответствующего вида», отмеченное ранее. Если я просто говорю: «Слово «лошадь» референциально указывает на объекты, обладающие свойством, появление которого заставляет меня при определенных обстоятельствах произносить предложение «Передо мной находится лошадь»», то одно из затруднений состоит в том, что существует слишком много подобного рода свойств. Например, пусть Л-Я (сокращение для «Лоша-

ди-Явления») будет тем свойством всех ситуаций восприятия, которое вызывает реакцию в виде произнесения предложения «Передо мной находится лошадь» обыкновенным человеком, владеющим английским языком. Тогда свойство Л-Я присутствует тогда, когда я говорю «Передо мной находится лошадь» (даже тогда, когда я испытываю оптический обман), однако слово «лошадь» референциально указывает скорее не на ситуации, обладающие этим свойством, а на определенную разновидность животных. Как присутствие животного, обладающего свойством принадлежности к отдельному естественному виду, так и присутствие ситуации восприятия, обладающей свойством Л-А связаны с произнесением мной высказывания «Передо мной находится лошадь» каузальными цепочками. В действительности, появление лошадей в каменном веке связано с моим высказыванием «Передо мной находится лошадь» каузальной цепочкой. Точно так же как существует слишком много сходств для того, чтобы референция была просто проблемой сходства, так и существует слишком много каузальных цепочек для того, чтобы референция была просто проблемой каузальных цепочек.

С другой стороны, если я говорю «Слово «лошадь» референциально указывает на объекты, обладающие свойством, которое связано с произнесением мною высказывания «Передо мной находится лошадь» при определенных обстоятельствах с помощью *каузальной цепочки определенного типа*», то передо мной встанет проблема, согласно которой в том случае, если я в состоянии определить, что собой представляет каузальная цепочка определенного типа, я уже должен быть способен референциально указывать на те виды вещей и свойств, которые образуют этот вид каузальной цепи. Однако как я оказался способным сделать это?

Заключение состоит не в том, что не существует терминов, которые обладали бы логикой, приписываемой теорией подобия, и не в том, что *не* существует референциально отсылающих к вещам терминов, которые связаны с нами при помощи различных каузальных цепочек. Заключение состоит просто-напросто в том, что ни подобие, ни каузальная взаимосвязь не могут быть ни единственными, ни фундаментальными механизмами референции.

ВИТГЕНШТЕЙН О «СЛЕДОВАНИИ ПРАВИЛУ»

Рассмотрим пример, который я мимоходом упомянул выше, т. е. пример с человеком, который пытается определить аспект R (т. е. аспект, в котором ощущения должны быть сходны с X , если они должны быть правильно классифицированы как E), говоря или думая, что две вещи сходны в отношении R только когда они подобны тем способом, каким подобны Z и W . Это, конечно же, не так, поскольку любые две вещи Z и W сходны в более чем одном отношении (в действительности, они сходны в бесконечном множестве отношений). Попытка определить отношение сходства при помощи задания конечного множества примеров подобна попытке определить функцию натуральных чисел путем перечисления первой тысячи или первого миллиона значений: всегда существует бесконечное множество функций, которые согласуются с любым данным списком конечного множества значений, но которые расходятся со значениями, не упомянутыми в этом списке.

Это обстоятельство связано с еще одним соображением, которое Витгенштейн выдвигает в «Философских исследованиях» и которое было упомянуто в конце первой главы. Какие бы интроспективно наблюдаемые знаки или «представления» я бы ни был в состоянии вызвать в связи с определенным понятием, они не могут определить или конституировать *содержание* этого понятия. Витгенштейн высказывает это соображение в знаменитом разделе, касающемся «следования правилу» — скажем, правилу «добавить единицу». Даже если два вида (*species*) в двух возможных мирах (я формулирую этот аргумент, *существенно* отклоняясь от той терминологии, которую использовал Витгенштейн!) обладают одними и теми же ментальными знаками в связи с вербальной формулой «добавить единицу», все-таки сохраняется возможность, что их *практика* могла бы расходиться; причем именно практика определяет интерпретацию: как мы видели, знаки не интерпретируют сами себя. Даже если кто-то изобразит отношение « A следует за B » (т. е. $A = B + 1$) точно так же, как это сделали мы, и выразит с нами согласие в отношении большого конечного множества случаев (например, что 2 следует за 1, что 3 следует за 2, ..., что 999,978 следует за 999,997), все же он может придерживаться иной интерпретации слов

«следовать за», что обнаружится только в будущих случаях. (Даже если он соглашается с нами в своей «теории» — т. е. в том, что он говорит относительно выражения «следовать за...», он, как показывает теорема Левенгейма-Сколема, может иметь отличающуюся от нашей интерпретацию всей теории в целом).

Это обстоятельство имеет самое непосредственное отношение как к философии математики, так и к философии языка. Прежде всего, здесь имеет место проблема *конечности*: человеческая практика, актуальная или потенциальная, всегда ограничена. Даже если мы утверждаем, что мы обладаем определенными способностями, мы все же неспособны «проводить перечисление до бесконечности». Если *существуют возможные различные расшифрения нашей практики, то существуют возможные различные интерпретации даже последовательности натуральных чисел* — наша практика, или наши ментальные репрезентации, и т. д. не выделяют какой-либо единственной «стандартной модели» последовательности натуральных чисел. Мы склонны думать, что они выделяют подобного рода модель потому, что легко переходим от высказывания «мы могли бы продолжать перечисление» к высказыванию «идеальная машина могла бы продолжать перечисление» (или, «идеальное сознание могло бы продолжать перечисление»); однако разговор об идеальных машинах (или сознаниях) очень сильно отличается от разговора о *реальных* машинах или личностях. Разговор о том, что могла бы делать идеальная машина, есть разговор, происходящий *в рамках* математики; он не может дать интерпретацию *самой* математике.

Точно так же Витгенштейн считает, что разговор о «сходстве» и о «том же самом ощущении» или о «том же самом опыте» есть разговор, происходящий *в рамках* психологической теории; он не может дать интерпретацию *самой* психологической теории. Итак, интерпретация психологической теории и терминологии определяется нашей реальной практикой, нашими реальными стандартами правильности и неправильности.

В «Способах создания миров»¹¹ Нельсон Гудмен делает сходное замечание: невозможно понять то, что собой представля-

11. Goodman N. The Ways of Worldmaking. Hackett Publishing Company, 1978. P. 92. [Ср.: Гудмен Н. Способы создания миров / Пер. с англ. М. В. Лебедева. М.: Идея-Пресс; Праксис, 2001. С. 207. — Прим. пер.].

ют «в действительности» факты восприятия независимо от способа их концептуализации, от описаний, которые мы им даем и которые представляются нам правильными. Таким образом, после обсуждения открытия психолога Колерса, согласно которому значительная часть инженеров и физиков вообще неспособна видеть явное движение, т. е. «движение», производимое светом, который последовательно вспыхивает в разных местах, Гудмен дает следующий комментарий:

«Тем не менее, если наблюдатель сообщает, что он видит две различные вспышки даже при настолько коротких расстояниях и интервалах, что большинство наблюдателей видят в этот момент одно перемещающееся цветное пятно, то, возможно, это означает, что он видит два пятна, поскольку мы могли бы сказать, что видим скопление молекул, когда смотрим на стул, или поскольку мы говорим, что видим круглую крышку стола даже тогда, когда мы смотрим на нее под наклонным углом. Так как наблюдатель может научиться хорошо отличать реальное движение от видимого, он может принимать видимое явление движения за знак, что произошли две вспышки, как мы принимаем видимое явление овальной крышки стола за знак, что она — круглая; и в обоих случаях знаки могут быть настолько прозрачными, что мы смотрим сквозь них на физические объекты и события. Когда наблюдатель визуально определяет, что то, что находится перед ним, — это то, что, как мы согласны, находится перед ним, мы едва ли можем обвинить его в ошибке визуального восприятия. Должны ли мы в таком случае утверждать, что он неправильно истолковывает предписание, которое требует от него, чтобы он сообщал исключительно о том, что он видит? Как в таком случае мы можем, не предпрещая заранее результата, переделать это предписание, дабы предотвратить подобного рода «недоразумение»? Попросить наблюдателя не пользоваться никаким предшествующим опытом и избегать всякой концептуализации — значит, очевидно, оставить его безмолвным, поскольку для того, чтобы вообще что-то высказать, он должен использовать слова».

ПОСТИЖЕНИЕ «ФОРМ» И ЭМПИРИЧЕСКАЯ АССОЦИАЦИЯ

Античный платоник или неоплатоник решил бы эту проблему более простым способом. Философ такого склада сказал бы, что когда мы следим за отдельным ощущением, мы воспринимаем также универсалию, или форму, т. е. что ум обладает способностью постигать свойства сами по себе, а не только отдельные случаи их проявления. Такой философ заявил бы, что именно номинализм Витгенштейна и Гудмена, их нежелание иметь какое-либо дело с формами и с непосредственным постижением форм, обусловил их мнение, что теория подобия порождает проблемы.

В то время как простое постулирование таинственной способности «постижения форм» вряд ли решит интересующую нас проблему, могло бы сложиться впечатление, что мы *в силах* обладать каким-то подобием этой способности. Свойства вещей и в самом деле входят в *каузальные объяснения*: когда я обладаю ощущением и это вызывает реакцию «это ощущение красного», моя реакция оказывается частично обусловленной тем фактом, что ощущение имеет *свойство*. Действительно, некоторые философы являются настолько радикальными номиналистами, что вообще стали бы отрицать существование таких сущностей, как «свойства»; однако сама наука, не колеблясь, вполне свободно рассуждает о свойствах. Не можем ли мы сказать, что когда изобретатель индивидуального языка Витгенштейна (т. е. человек, который хотел изобрести индивидуальный язык) наблюдал *Х* и говорил «*Е*», то реакция «*Е*» была обусловлена каузальным взаимодействием, включающем в себя определенное *свойство*, и что *это* свойство (каким бы оно ни было) и есть то «подобие», в отношении которого другие ощущения должны находится к ощущению *Х* для того, чтобы классифицироваться как *Е*?

Замечание, согласно которому разговор о «свойствах» с научной точки зрения является совершенно правомочным, правильно; однако оно вовсе не помогает нам реабилитировать платонизм. Мы взаимодействуем со свойствами только путем взаимодействия с *частными случаями* их проявления; и эти частные случаи всегда являются частными случаями проявления *множества* свойств в одно и то же время. Не существует такой вещи, как взаимодействие со свойством «самим по себе». Разговор о

свойствах, каузально ассоциированных с ощущением, не может выполнять той функции, которую понятие (единственной в своем роде) формы ощущения выполняет в платоновской философии.

Раскроем смысл вышеприведенного замечания: когда я имею ощущение голубого цвета, я имею ощущение *голубого*, а также я имею ощущение, обладающее сложным свойством «быть таким, чтобы классифицироваться мной в данном случае под данной вербальной этикеткой». Простое наблюдение за *данным* ощущением не конституирует «постижения» *одного* из этих свойств. Выделение свойства, ассоциированного как раз *одним* из этих способов с моим ощущением или с вербальной этикеткой — это наш старый приятель, т. е. опять-таки проблема каузальной цепи соответствующего типа.

Для того чтобы понять это, отметьте, прежде всего, что когда мой совокупный перцептивный опыт вызывает реакцию «я обладаю ощущением голубого цвета», я не всегда *прав*. Мне самому приходилось дважды или трижды указывать на «человека в голубом свитере», прежде чем кто-то не делал замечание, что свитер является *зеленым*. Я не имею в виду то, что свитер казался *голубым*; я понял, что я неправильно описывал свитер, о котором говорил другой человек. (Мне не часто доводилось говорить «я имею ощущение голубого цвета», но если я делал это, то в этом случае я говорил это дважды или трижды, прежде чем кто-то, удивленный, вероятно, тем, почему это я имею ощущение голубого цвета, в то время как смотрю на что-то, что явно имеет зеленый цвет, — не поправлял меня, вследствие чего я брал назад мой прежний феноменальный отчет). Это показывает, что свойство вызывать отчет «я имею ощущение голубого цвета», или какой угодно еще, *не является тем же самым свойством*, что и свойство быть ощущением голубого цвета, или ощущение чего бы то ни было еще, чем могло бы быть соответствующее качество.

Философы часто ссылаются на «оговорки». Суждения подобного рода представляются мне неудачными. Я мог бы намереваться произнести слово «зеленый», но оказалось, что я случайно *произнес* слово «голубой». Это было бы оговоркой. Однако в описанном мною случае я даже не замечал, что мое описание ошибочно, до тех пор, пока кто-то не подверг мой отчет сомне-

нию (в противном случае я вообще мог бы не заметить своей ошибки).

Другое объяснение, которое приходит на ум, заключается в том, что когда я произносил «голубое», я *имел в виду* зеленое. На данный момент должно быть ясно, что когда мы говорим о вещах, мы не имеем дела со «значениями» вещей в том смысле, что мы имеем эти значения в уме. Сказать я «имел в виду» зеленое означает просто сказать, что я незамедлительно принял поправку (и посмеялся, когда понял, что я говорил). Это означает просто повторить то, что произошло, а не объяснить его.

Каким бы ни было объяснение (возможно, какой-то сбой в той части моего мозга, которая отвечает за речь), суть проблемы заключается в том, что точно так же, как свойство Л-Я, описанное несколькими страницами выше, будет вызывать отчет «передо мной находится лошадь» даже в тех случаях, когда в окружающей среде не будет присутствовать ни одной лошади, так и в нашем последнем случае имеется сложное свойство моего общего умонастроения, которое будет вызывать произнесение высказывания «я имею ощущение голубого цвета» в то время, когда я не имею ощущения голубого цвета (или, во всяком случае, отрицало бы, что я имею такое ощущение, если бы меня об этом спросили). Ни один механизм эмпирической ассоциации не является совершенным. Если вы решите оговорить в качестве особого условия, что я имею ощущение голубого цвета всякий раз, когда я имею ощущение, которое *вызывает* этот отчет (или которое вызывает этот отчет и является таковым, что отчет не кажется мне «ошибочным» при повторном размышлении), то согласно теории народной психологии и, возможно, согласно также и научной психологической теории, могли бы существовать обстоятельства, при которых было бы истинно, что я имею ощущение голубого цвета согласно *этому* критерию, хотя, в силу одного из множества оснований, качество ощущения не является голубым. Более того, как отмечает Витгенштейн, согласно этому критерию *все то, что кажется мне правильным, оказывается правильным*, — т. е. различие между отчетом о моем ощущении, который является правильным, и отчетом о моем ощущении, который кажется мне правильным, будет отброшено. Возможно, нам *следовало бы* отбросить или по крайней мере уточнить его; возможно, как, по-видимому, полагает Гудмен,

вопрос о том, обладает ли человек «реально» той разновидностью ощущения, которой он, как сам думает, обладает, не имеет смысла, кроме как в особых случаях; таких, когда человек берет свой отчет назад после того, как он подвергается сомнению; однако для метафизического реалиста отрицание этого различия является невозможным шагом, поскольку резкое различие между тем, что имеет место, и тем, имеет место по чьему-то мнению и есть тот фактор, что конституирует метафизический реализм.

МОЖНО ЛИ ПОСТОЯННО ЗАБЛУЖДАТЬСЯ ОТНОСИТЕЛЬНО КАЧЕСТВ СВОИХ ПРОШЛЫХ ОЩУЩЕНИЙ?

Другой способ выяснить то, о чем идет речь, — это рассмотреть вопрос: «Можно ли постоянно заблуждаться относительно своих прошлых ощущений?» Согласно теории подобия, ответ, конечно же, положительный; ибо согласно этой теории, мои прошлые ощущения или сходны, или не сходны с ощущениями, которые я *в настоящее время* описываю при помощи различных вербальных этикеток типа «ощущение красного», «боль» и т. д., и вопрос о сходстве или несходстве совершенно отличается от вопроса о том, классифицировал ли я их *тогда* при помощи тех же самых вербальных этикеток. Возможно, мир устроен так, что то, что мы называем «ощущением красного», в некоторое четное мгновение от начала нашей эры является актуально сходным по своему качеству с тем, что мы называем «ощущением зеленого» в некоторое нечетное мгновение, однако наша память постоянно вводит нас в заблуждение таким образом, что мы никогда этого не замечаем. В таком случае ощущение, которое я мгновение назад классифицировал под вербальной этикеткой «ощущение красного», *не* будет похоже на ощущение, которое я в настоящий момент классифицирую под той же самой этикеткой.

В этой утверждаемой возможности есть, однако же, что-то весьма странное. Прежде всего, довольно странно выглядит смысл фразы «я бы никогда не заметил»: если я истолковываю «свои ощущения красного» в различные моменты времени как заслуживающие доверия знаки различных соотнесенных физических событий (таких, как огонь, сигнал светофора и т. д.), то

мне будет сопутствовать успех во всех моих действиях. «Неправильный» класс сходства (класс, который смешивает в одну кучу ощущения, которые я *называю* ощущениями красного, несмотря на тот факт, что они «реально» не обладают одним и тем же «качеством») был бы таким классом, который мне *лучше* было бы использовать в связи с моей деятельностью по решению проблем. Но является ли он в этом случае классом *неправильного* сходства?

Если мы не предполагаем, что понятие сходства — это такое понятие, которое само дает себе интерпретацию, то тогда этот случай можно было бы переописать как такой случай, в котором отношение, называемое «сходством» внешним наблюдателем, рассказывающим нам об этом случае, просто отличается от отношения, которые мы называем «сходством». Если мы примем этот взгляд, то «гипотеза», согласно которой мы «реально» ошибаемся в отношении наших прошлых ощущений, рухнет: с *интерналистской* точки зрения не существует какого-либо вразумительного понятия ощущений, которые были бы «сходны» друг с другом в различные моменты времени безотносительно к нашим стандартам рациональной приемлемости.

СНОВА ТЕОРИЯ ИСТИНЫ КАК СООТВЕТСТВИЯ

Теперь читатель может убедиться в том, что теория референции как подобия является целиком и полностью ошибочной. Однако почему нам следовало бы сделать вывод, что теория истины как соответствия должна быть отброшена? Даже если понятие «сходства» между нашими понятиями и тем, на что они референциально указывают, не работает, то разве не могла бы существовать какая-то разновидность абстрактного изоморфизма, или, если не буквально изоморфизма, то какая-то разновидность абстрактного *отображения* понятиями вещей, находящихся в (независимом от сознания) мире? Не могла ли бы истина быть определена в терминах такого изоморфизма или отображения?

Проблема, связанная с этим предположением, состоит не в том, что не существует соответствий между словами или понятиями и другими сущностями, но в том, что существует *слишком много* таких соответствий. Для того, чтобы выделить только *одно*

соответствие между словами или ментальными знаками и вещами, независимыми от сознания, мы должны были бы уже иметь референциальный доступ к независимым от сознания вещам. Вы не можете выделить соответствие между двумя вещами, сильно сжав только *одну* из них (или сделав что-то еще с одной из них); вы не можете выделить соответствие между нашими понятиями и предполагаемыми ноуменальными объектами, не имея доступа к ноуменальным объектам.

Понять это, к примеру, можно следующим образом. Иногда противоречащие друг другу теории могут быть взаимно переводимыми. Например, если бы физика Ньютона была истинной, то каждое отдельное физическое событие можно было бы описать двумя способами: в терминах частиц, действующих на расстоянии в пустом пространстве (именно так Ньютон описывает действие гравитации), или же в терминах частиц, воздействующих на поля, которые воздействуют на другие поля (или же на другие части того же самого поля), которые в конечном счете воздействуют «локально» на другие частицы. Например, уравнения Максвелла, которые описывают поведение электромагнитного поля, математически эквивалентны теории, согласно которой между частицами существуют только силы, действующие на расстоянии, притягивающие и отталкивающие по закону обратной пропорциональности квадрату расстояния, распространяющиеся не мгновенно, но скорее со скоростью света («замедленные потенциалы»). Теория поля Максвелла и теория замедленного потенциала с метафизической точки зрения противоречат друг другу, поскольку либо существуют, либо не существуют каузальные факторы («поля»), которые опосредуют действия отдельных частиц друг на друга (как сказал бы реалист). Однако эти две теории математически взаимопереводимы. Так что если имеет место «соответствие» ноуменальным вещам, которое делает одну из этих теорий истинной, то в этом случае можно определить другое соответствие, которое делает другую теорию истинной. Если все, благодаря чему теория становится истинной, — это абстрактное соответствие (неважно, какого рода), то противоречащие друг другу теории обе могут быть истинными.

Для интерналиста в этом выводе нет ничего, что заслуживало бы возражения: почему бы не существовать равно согласован-

ным, но противоречащим друг другу концептуальным схемам, которые в равной степени удовлетворяют нашим почерпнутым из опыта убеждениям? Если истина не является (одним-единственным) соответствием, то тогда открывается возможность определенного рода плюрализма. Однако метафизический реалист движим желанием спасти понятие Точки Зрения Божественного Взора, т. е. Одной Истинной Теории.

Соответствие может существовать не только между объектами и тем, что мы считаем несовместимыми теориями (т. е. *те же самые объекты* могут быть, как принято говорить у логиков, «моделью» для противоречащих друг другу теорий); даже если мы зафиксируем теорию и зафиксируем объекты, то все равно существует (в том случае, если число объектов бесконечно) бесконечно много *различных* способов, которыми те же самые объекты могут использоваться для того, чтобы создать модель для данной теории. Это соображение просто выражает на математическом языке интуитивный факт, согласно которому для того, чтобы установить соответствие между двумя областями, необходимо обладать независимым доступом к обеим областям.

Мы присутствуем при кончине теории, просуществовавшей почти что две тысячи лет. Тем, что эта теория просуществовала столь долго и принимала столь разнообразные формы, несмотря на свойственные ей с самого начала внутренние противоречия и несообразности, она обязана естественности и силе желания занять Точку Зрения Божественного Взора. Кант, который первый учил нас тому, что это желание невыполнимо, тем не менее считал, что оно встроено в саму нашу рациональную природу (он предполагал, что можно воплотить этот «всеохватывающий» импульс в проекте, направленном на осуществление «высшего блага в мире» путем согласования морального и эмпирического порядков в совершенной системе социальных учреждений и личных отношений). Продолжающееся присутствие этого естественного, но невыполнимого импульса является, по всей вероятности, фундаментальной причиной ложных монизмов и ложных дуализмов, расплодившихся в нашей культуре; но как бы то ни было, мы отказываемся от Точки Зрения Божественного Взора.

4. СОЗНАНИЕ И ТЕЛО

ПАРАЛЛЕЛИЗМ, ИНТЕРАКЦИОНИЗМ, ТОЖДЕСТВО

В XVII столетии все величайшие философы — Декарт, Спиноза и Лейбниц — отдавали себе отчет в том, что существует серьезная проблема, связанная с отношением между сознанием и материальным телом. В известном смысле это отношение представляло собой проблему уже для Платона, равно как и для всех философов, живших и творивших после него; однако оно стало еще более проблематичным после появления современной *физики*. В XVII столетии люди осознали, что материальный мир является *по сути своей каузально замкнутым* миром. Тот способ, благодаря которому мир оказывается каузально замкнутым, лучше всего поддается выражению в терминах физика Ньютона: тело движется лишь постольку, поскольку на него воздействует какая-то *сила*. Силы поддаются исчерпывающему описанию при помощи чисел: трех чисел достаточно для того, чтобы определить направление, и одного числа достаточно для того, чтобы определить величину силы. Ускорение, инициированное силой, имеет в точности то же самое направление, что и сила, и величина ускорения может быть дедуцирована из массы тела и величины силы в соответствии с первым законом Ньютона — $F = ma$. Когда тело испытывает воздействие более чем одной силы, суммарная сила может быть вычислена в соответствии с законом параллелограмма.

Необходимо отдавать себе отчет в том, насколько эта физика, подчеркивающая значение числа и точных алгоритмов вычисления, отличается от по существу своему *количественного* мышления средних веков. В средневековом мышлении практически любая вещь была способна «повлиять» на что угодно. (Наше нынешнее словечко «инфлюэнца» представляет собой пережиток средневекового склада мышления. Было принято считать, что злые духи — по-итальянски *questa influentia* — оказывают влияние на воздух, который, в свою очередь, заражает людей). Если придерживаются подобного склада мышления, то нет ничего удивительного в том, что сознание может «воздействовать» на тело.

В Новое время, когда творили философы, о которых я упомянул выше, появляется математический образ мышления; он стал постепенно вытеснять старый способ мышления. Свою завершенную форму математический образ мышления приобрел только в произведениях Ньютона, однако уже у Декарта можно в ряде случаев встретить упоминания о параллелограмме сил, более простые примеры, связанные с параллелограммом, можно найти уже у Леонардо да Винчи. Эти мыслители уже предвосхищали будущий образ физики. Они прекрасно отдавали себе отчет в том, что физика имеет дело с силой и движением, поэтому они отказались от выдвижения объяснений, сформулированных при помощи терминов, обозначающих качества. Скорее они исходили из того, что механистический мир имеет свою собственную логику, имеет свою собственную «программу», как сказали бы мы, и мир следует этой программе до тех пор, пока что-то не разрушит его.

Эти мыслители считали, что ментальные события могут существовать одним из двух различных способов. (1) Они могли бы происходить *параллельно* физическим событиям, то есть событиям, локализованным в человеческом мозге. Эта модель представляет собой пару часов, показывающих одинаковое время: тело — это часы, которые были заведены и которые идут своим ходом вплоть до смерти; соответственно, физический мир в целом — это тоже часы, которые идут своим ходом от сотворения мира и вплоть до судного дня (или же, если исходить из современных представлений, вплоть до гравитационного коллапса вселенной). Ментальные события тоже идут своим чередом, и каким-то образом, быть может, благодаря божественному проведению, было устроено так, что событие *B* в мозге всегда происходит только тогда, когда имеет место ощущение *S*. (2) Они могли бы *взаимодействовать* с физическими событиями. Ментальные события могли бы на деле вызывать события в мозге, и наоборот.

Печально известная форма интеракционизма, предложенная Декартом, сводящаяся к предположению, что сознание может оказывать воздействие на материю, тогда как материя является чем-то весьма и весьма эфемерным и что сознание в действительности каким-то образом воздействует на материю через шишковидную железу, была значительно менее безумной, неже-

ли она могла бы показаться на первый взгляд. По сути дела, она представляла собой пережиток ряда средневековых учений¹. Согласно более ранним представлениям, сознание воздействует на «дух», который, в свою очередь, воздействует на «материю», при этом дух вовсе не считался *исключительно* нематериальной сущностью. «Дух» относился к той опосредующей разновидности материала, который средневековых философов заставляла постулировать их склонность вводить посредников между любыми двумя смежными терминами в ряду видов бытия. «Дух» чем-то напоминал газ, обладающий очень невысокой атмосферой давления. Как только «дух» был отброшен, а сознание стали мыслить как совершенно нематериальную сущность, воздействие сознания пусть даже на весьма эфемерную материю в виде шишковидной железы стало казаться довольно-таки странным. Подобного рода взаимодействие невозможно было себе мысленно представить.

Наиболее бесхитростная версия интеракционизма понимает сознание как своего рода призрак, способный вселяться в различные тела (при этом его способность мыслить, чувствовать, предаваться воспоминаниям и демонстрировать признаки тождества личности остается неизменной, в противоположность мнениям, заимствованным из потока популярных книг, посвященных перевоплощению и «воспоминаниям о предшествующих жизнях») или даже способный существовать без тела, продолжая при этом мыслить, чувствовать, вспоминать, и демонстрировать признаки тождества личности. Эта версия, которая, по сути дела, представляет собой не что иное, как суеверие, уязвима для возражения, согласно которому существует громадное количество данных, некоторые из которых были известны уже в XVII веке, свидетельствующих о том, что исполнение функций мышления, чувств и памяти в обязательном порядке требует участия *мозга*. Конечно, из этой версии совершенно непонятно, почему мы вообще должны были бы обладать сложно организованным мозгом. Если все, что нам требуется — это «рулевое колесо», то оно могло бы быть значительно меньше, чем человеческий мозг.

1. По поводу описания средневековой точки зрения см.: Lewis C. S. *The Discarded Image*. Cambridge, 1964. Chapter VII, sec. F.

Для того, чтобы избежать подобного рода научных возражений, изошренные интеракционисты, подобно Декарту, считали, что сознание и мозг представляют собой *сущностное единство*. Именно сознание и мозг в их единстве выступают в роли той действующей силы, которая мыслит, чувствует, вспоминает и демонстрирует признаки тождества личности. Это означает, что сущность, которую мы обычно называем сознанием, это вовсе не сознание, а сознание и мозг в их единстве. Не вполне понятно, однако, что же, собственно говоря, означает это учение, то есть что означает утверждение, что какая-то сущность может состоять из двух субстанций, столь же отличающихся друг от друга, сколь, согласно общепринятым представлениям, отличаются друг от друга сознание и материя, но в то же самое время представлять собой сущностное единство.

Альтернатива, выдвигаемая параллелизмом, также выглядит весьма странно. Благодаря чему ментальное событие сопровождает событие, происходящее в мозге? Один смелый мыслитель XVII века предположил, что ментальные события могли бы в действительности быть *тождественны* событиям, происходящим в мозге, и иным физическим событиям. Этим мыслителем был Спиноза. Это предположение, если сформулировать его в современной форме, утверждает, что событие, выражающееся в том, что я при определенных обстоятельствах испытываю чувство боли, могло бы быть тем же самым событием, которое происходит в моем мозге, пребывающем в некотором состоянии *В* при тех же самых обстоятельствах. (Я бы выразил эту же точку зрения, сказав, что свойства обладания определенным ощущением боли и пребывания моего мозга в некотором состоянии *В* являются тождественными. Я отдаю предпочтение именно второй формулировке потому, что, на мой взгляд, в настоящее время логическая теория *свойств* разработана значительно лучше, чем логическая теория событий; сама идея, однако же, может быть изложена обоими способами. В соответствии с предложенной мною терминологией, идея состоит в том, что свойство личности, а именно то, что личность испытывает ощущение *Q*, могло бы быть *тем же самым свойством*, что и свойство мозга пребывать в состоянии *В*). В этой форме данное предположение было выдвинуто, к примеру, Дидро в XVIII столетии; оно заняло господствующие позиции в 40-х и 50-х годах XX века. Материализм

и теория тождества впервые стали восприниматься всерьез, и стало выдвигаться предположение, что правильным является взгляд, подобный взгляду Спинозы (или, точнее, — взгляд, разделявшийся Спинозой, но *освобожденный* при этом от его тщательно разработанных теологических и метафизических преувеличений). Суть этого взгляда сводится к следующему: мы действительно имеем дело с одним миром, и то обстоятельство, что до тех пор, пока мы не проведем основательные научные исследования, нам остается неизвестным, что ощущения боли, обладание слуховыми и зрительными ощущениями — это на самом деле не что иное, как состояния мозга, вовсе не доказывает того, что подобного рода тождества не существует.

Первая современная форма теории тождества была выдвинута несколькими мыслителями, наиболее известным из числа которых является австралийский философ Дж. Д. Сمارт. Первоначально предположение сводилось к тому, что ощущение, к примеру, отдельное ощущение голубого цвета, тождественно определенному нейрофизиологическому состоянию. Одной из разновидностей этой точки зрения, выдвинутой первоначально мною самим, была теория, названная *функционализмом*². Согласно функционалистской точке зрения, подобного рода тождество между ментальными и нейрофизиологическими состояниями действительно имеет место, однако Смарт давал неправильную характеристику состоянию мозга как составной части этого тождества. Функционалист считает, что мозг обладает свойствами, которые в известном смысле *не являются* физическими.

Итак, что же собственно говоря, я имею в виду, когда рассуждаю о том, что *мозг имеет не-физические* свойства? Я имею в виду такие свойства, которые *поддаются определению при помощи терминов, не содержащих ссылок на физический или химический состав мозга*. Если вам кажется странным, что физическая система обладает свойствами, которые не являются физическими, давай-

2. Выпущенная под редакцией Нэда Блока хрестоматия (*Block N. (ed.). Readings in Philosophy of Psychology*. Harvard, 1980) содержит прекрасную подборку статей, посвященных функционализму. Мои собственные статьи по этой теме опубликованы в качестве глав 14—22 моей книги «Сознание, язык и реальность». См.: *Putnam H. Mind, Language and Reality. Philosophical Papers. Vol. 2*. Cambridge, 1975.

те рассмотрим вычислительную машину. Вычислительная машина имеет самые разнообразные физические качества. Например, она обладает определенным весом; она содержит массу микросхем, по которым циркулируют сигналы, и т. д. Она обладает также экономическими свойствами, например, определенной ценой. Кроме того, она имеет еще и функциональные свойства: например, в нее заложена определенная программа. Свойство последнего типа является не-физическим *в том смысле, что оно может быть реализовано системой совершенно независимо от того, что собой представляет ее метафизическое или же онтологическое строение*. Бестелесный дух мог бы действовать по определенной программе, мозг мог бы действовать по определенной программе, машина могла бы действовать по определенной программе, и функциональная организация этих трех сущностей — бестелесного духа, мозга и машины — могла быть полностью одинаковой несмотря на то, что материал, из которого они состоят, целиком и полностью различен.

Психологические свойства проявляют те же самые характеристики; то же самое психологическое свойство (например, быть рассерженным) может быть свойством существ, принадлежащих к тысячам самых разнообразных видов, которые могут обладать совершенно различным физическим строением и химическим составом; некоторые из этих существ могут быть неземными существами; возможно, бывает так, что даже роботы сердятся. Предположение функционализма заключается в том, что наиболее убедительная «монистическая» теория XX века, теория, которая позволяет не рассматривать сознание и материю как два различные типа субстанций или же как два отдельные царства бытия, утверждает, что психические свойства тождественны функциональным свойствам.

В настоящее время я всё еще склонен полагать, что эта теория является правильной; по крайней мере я убежден, что она представляет собой правильное *натуралистическое* описание отношения между сознанием и телом. Существуют иные, «менталистские» описания этого отношения, которые то же являются правильными, но которые в то же самое время не сводимы к той картине мира, которую мы именуем «Природой» (само собою разумеется, что понятия «рациональность», «истина» и «референция» *относятся* к подобной «менталистской» версии). Об

этом речь у меня пойдет несколько позже (см. главу 6). Этот факт меня вовсе не смущает: ведь, как подчеркнул Нельсон Гудмен, одна из привлекательных особенностей не-реалистической позиции состоит в том, что она допускает возможность существования альтернативных правильных версий мира. Меня, однако же, привлекает идея, согласно которой *единственной* правильной версией является натуралистическая версия, в рамках которой формы мышления, образы, ощущения и т. д. *представляют собой* функционально характеризуемые физические события. В данном же случае я намереваюсь обсудить одно связанное с функционалистской теорией затруднение, которое пришло мне в голову несколько лет назад. Суть этого затруднения сводится к тому, что функционалистская теория сталкивается с проблемами при объяснении качественного характера ощущений. Когда речь идет об относительно абстрактных, чистых психологических состояниях, например о том, что мы называем убеждением, «заключенным в скобки», то есть о мысли, которая рассматривается исключительно в ее «понятийном» содержании, или же о таких распространенных эмоциональных состояниях, как состояние ревности или же состояние раздраженности, то идентификация этих состояний с функциональными состояниями всей системы в целом представляется весьма убедительной. Но когда речь заходит о том, что человек обладает каким-то данным ему качеством, например, ощущает определенный оттенок голубого цвета, идентификация выглядит неправдоподобной.

Много лет я использую в лекциях пример, который представляет собой вариант известного примера «перевернутого спектра». Пример перевернутого спектра (появляющийся у Локка³⁾) рассказывает о человеке, который видит вещи так, что синее для него выглядит красным, а красное синим (или так, чтобы его субъективные цвета выглядели скорее негативом, чем позитивом — черное белым, белое черным). Первая реакция на такой случай могла бы быть следующей: «Бедный парень, люди должны пожалеть его». Но откуда мы это знали бы? Когда он видит что-нибудь синее, это выглядит для него красным, но он обучен назвать этот цвет *синим* с тех пор, как он был младен-

3. Locke J. An Essay Concerning Human Understanding. Book II. Chapter 32 (sec. 14).

цем, так что если его спросят, какого цвета этот объект, он ответил бы «синий». Так что никто никогда ни о чем и не узнал бы.

Мой вариант этого примера таков: вообразите, что однажды в вашей жизни ваш спектр становится перевернутым в обратную сторону, и вы *помните то, каким он был раньше*. Здесь нет никакой эпистемологической проблемы «проверки». Вы просыпаетесь однажды утром, и небо выглядит красным, а ваш красный свитер, оказывается, стал синим, а все лица — ужасного цвета, как на фотонегативе. Какой кошмар! Но теперь, возможно, вы сможете научиться поменять ваш способ речи и называть вещи, которые выглядят для вас красными — «синими», и, возможно, вы настолько в этом преуспеете, чтобы, если кто-то спросит вас, какого цвета его свитер, то вы дадите «нормальный» ответ. Но по ночам, предположим, вы будете стонать: «О, как мне жаль, что цвета больше не выглядят так, как когда я был маленьким. Цвета больше не выглядят так, как раньше».

В этом случае кажется, что мы даже знаем, что, по всей вероятности, случилось. Некоторые «провода» в мозгу, должно быть, «пересеклись». Входы от синего света, которые раньше шли в один механизм в мозгу, теперь идут в другой, а входы от красного света идут в первый. Другими словами, что-то изменилось в *реализациях*, в *физических* состояниях. *Физическое* состояние, которое прежде играло *функциональную* роль передачи сигналов присутствия «объективно» синего в окружающей среде, теперь сигнализирует о присутствии в ней «объективно» красного.

Теперь предположим, что мы приняли следующую «функционалистскую» теорию субъективного цвета: «некоторое ощущение является ощущением синего цвета (то есть имеет *качественный характер*, который я *сейчас* описываю этим способом) в том случае, если ощущение (или соответствующее физическое событие в мозгу) выполняет роль передачи сигналов присутствия объективно синего цвета в окружающей среде». Эта теория ухватывает один из смыслов выражения «ощущение синего», но не искомый нами «качественный» смысл. Если эта функциональная роль была бы *идентична* с качественным характером, то нельзя было бы сказать, что качество ощущения изменилось. (Если это не ясно, то представьте себе, что после инверсии спектра, и после того, как вы обучились лингвистической компенсации — называть красное синим и т. п., вы испытываете при-

ступ амнезии, которая стирает всю память о том, как цвета выглядели раньше. В этом случае казалось бы, что ощущение, которое вы *сейчас* называете «ощущением синего», имеет сейчас почти точно ту же функциональную роль, что имело раньше то ощущение, которое вы раньше называли «ощущением синего» — при том, что они полностью различны по характеру.) Но качество *изменилось*. Качество, как мне представляется, в этом случае не является *функциональным* состоянием.

Я думаю, что наиболее вероятный ход для функционалиста, если такие случаи действительно возможны — сказать: «Да, но «качественный характер» является только физической реализацией». И сказать, что для этого специального вида психологического свойства, *для качеств*, правильной была бы более старая форма теории идентичности. Если читатель ориентирован скорее материалистически, то он, вероятно, думает, что свойство наличия ощущения — это мозговое свойство. Читатели без материалистических склонностей, вероятно, думают, что свойство наличия ощущения *коррелирует* с мозговым состоянием. Вероятно, большинство людей поддержит один из этих двух подходов: либо подход, согласно которому состояния ощущения коррелируют с мозговыми состояниями, либо подход, согласно которому состояния ощущения идентичны с мозговыми состояниями. Как это часто случается, вопрос обсуждается вновь и вновь одним и тем же образом. Вот как он всегда обсуждается: «Если дано, что *B* коррелирует с *Q*, является ли *B* действительно идентичным с *Q*?» Мы знаем, что это состояние ощущения параллельно с этим мозговым состоянием; обстоит ли или нет дело так, что «состояние ощущения является идентичным с мозговым состоянием»? Чем дальше обсуждение продолжается этим способом, тем больше понятие *корреляции* начинает казаться непроблематичным. *Корреляция* не (очень) обсуждается, потому что *каждый знает*, что здесь наличествует по крайней мере корреляция. *Идентичность* обсуждается, потому что именно это и проблематично. Но я попробую показать, что даже *корреляция* является проблематичной — не в том смысле, что у нас здесь есть очевидность *отсутствия корреляции*, но в эпистемологическом смысле: *если* здесь наличествует корреляция, то мы не можем знать, какова она. Проблема будет зависеть не от того, чтобы принять материалистическую точку зрения, но от того

факта, что мы способны помыслить здесь по меньшей мере корреляцию.

ТЕОРИЯ ТОЖДЕСТВА И A PRIORI

Возрождение интереса к теории тождества и другим «монистическим» теориям, имевшее место в 60-е годы, стало возможным (если не принимать во внимание Смарта и других ранних сторонников теории тождества), благодаря изменению эпистемологического климата. Вплоть до 60-х годов философы не принимали всерьез теорию тождества потому, что «считали» ее ложной. Они полагали, что им доподлинно «известно», что эта теория является ложной, причем не на основании эмпирических данных (поскольку какого рода эмпирические данные могут показать, что ощущение не является состоянием мозга?), а *a priori*. Человек размышляет об этой теории и, руководствуясь априорными соображениями, видит, что ощущение не могло бы быть состоянием мозга. Или же, возможно, человек понимает, что говорить о том, что ощущение — это состояние мозга, столь же бессмысленно, как бессмысленно говорить о том, что число три является голубым. Вплоть до начала 50-х или даже 60-х годов XX века люди думали, что они *отлично знают*, что ощущения *не могут* быть физическими состояниями; многие думали, что они по крайней мере знают это. Другие же полагали, что подобного рода точка зрения является ошибочной. Однако доказать что-либо кому-либо было невозможно. Большинство сказало бы: «Да, мы не можем доказать вам, что ощущение не является нейрофизиологическим состоянием; да, мы не можем доказать вам, что за одним числом всегда следует другое; да, мы не можем доказать вам, что число три не является голубым; однако все это — истины, которые *не нуждаются в доказательстве*, иными словами, это — *истины разума*. Нам прекрасно известно, что ощущение не может быть нейрофизиологическим состоянием; утверждать нечто подобное было бы полной бессмыслицей». Итак, существовало большинство, которое знало, что ощущения не могут быть состояниями мозга, и меньшинство, которое знало, что большинство заблуждается. Каждая из полемизирующих сторон была уверена в том, что противоположная сторона заблуждается, причем эта уверенность основывалась исключительно на

априорных соображениях. Не было никакой возможности привести какое-нибудь доказательство в защиту своей позиции или же перейти от этого инертного состояния к дискуссии.

В 1951 году У. В. Куайн опубликовал статью, озаглавленную «Две догмы эмпиризма»⁴. С той поры доверие философов к понятию априорной истины постоянно уменьшалось. Куайн указал, что множество истин, которые мы считали известными нам *a priori*, оказалось нуждающимся в исправлении. Так, рассмотрим следующий пример: предположим, что кто-нибудь сказал бы Евклиду, что могут существовать две *прямые*, перпендикулярные *третьей*, которые при этом *пересекались бы* друг с другом. Евклид бы на это ответил, что этого не могло бы быть; знание о том, что две параллельные прямые не могут пересекаться друг с другом — это необходимая истина. Однако согласно физической теории, которой мы придерживаемся в настоящее время, подобного рода прямые *и в самом деле* пересекаются. Свет, распространяющийся неподалеку от солнца, ведет себя так не потому, что свет распространяется по кривым линиям, а потому, что свет продолжает распространяться по прямым линиям и прямые линии ведут себя подобного рода образом в нашем не-евклидовом мире.

Коль скоро мы соглашаемся с *этим*, то некоторые философы были вынуждены задать вопрос: «Что же остается от *a priori*?». Собственно говоря, Куайн так и поступает. (Кроме того, он довольно убедительно показывает, что стандартные эмпиристские объяснения *a priori*, например, при помощи понятия «истинности в силу конвенции», были непоследовательными и противоречивыми, однако я не буду здесь рассматривать его аргументы).

На мой взгляд, кое в чем Куайн заходит слишком далеко. Утверждение Куайна, что «ни одно высказывание не гарантировано от исправления» подразумевает, что для всякого высказывания существуют такие условия, при котором будут существовать *рациональные* мотивы для того, чтобы его отвергнуть. Однако достаточно очевидно, что это утверждение является ложным; к примеру, при каких обстоятельствах существовали бы

4. Статья «Две догмы эмпиризма» впервые появилась в 1951 году в журнале «Philosophical Review». Она была перепечатана в сборнике статей Куайна, озаглавленном «С логической точки зрения». См.: Quine W. V. Two Dogmas of Empiricism // Quine W. V. From the Logical Point of View. N. Y., 1961.

рациональные основания для того, чтобы отвергнуть высказывание «Не все высказывания являются истинными», то есть для того, чтобы принять высказывание «Все высказывания являются истинными»⁵?

Хотя Куайн, критикуя понятие *a priori*, допускает известного рода преувеличения, он все равно прав в главном: наши понятия рациональности и рациональной исправимости (*revisability*) не фиксируются при помощи какого-либо неизменного набора правил, точно также как не являются они присущими, как полагал Кант, нашей трансцендентальной природе; дело в том, что существуют весьма убедительные соображения, согласно которым сама идея трансцендентальной природы, то есть природы, присущей нам *ноуменально*, и независимой от объяснения наших личностей при помощи биологических и исторических терминов, лишена какого бы то ни было смысла. Поскольку наши понятия рациональности и рациональной исправимости являются продуктом нашего сильно ограниченного опыта и нашей склонной к ошибкам биологической конституции, необходимо ожидать того, что даже принципы, которые мы считаем «априорными» или «концептуальными», будут время от времени оказываться нуждающимися в пересмотре в свете неожиданного опыта или же непредусмотренных теоретических нововведений. Подобного рода исправление не может быть неограниченным; если бы оно носило неограниченный характер, то мы не имели бы никакого понятия о том, что можно было бы назвать «рациональностью»; однако мы не в состоянии установить в общем и целом подобные границы. За исключением тривиальных случаев, например, высказывания «Не все высказывания являются истинными», мы не можем быть уверенными в том, что у нас *никогда* не возникнет рациональных оснований для того, чтобы в *каком-нибудь* контексте отбросить высказывание, которое считается (причем, если брать *данный* контекст, вполне правомочно), «необходимой» истиной. В целом нам следует признать, что соображения полезности, всеобъемлющей полезности и правдоподобия могут заставить нас отбросить истину,

5. Я обсуждаю критику Куайном понятия *a priori* в статье «Аналитичность и априорность». См.: Putnam H. Analyticity and Apriority: Beyond Wittgenstein and Quine // Midwest Studies in Philosophy. Vol. 4. Minnesota, 1979.

считавшуюся прежде априорной, и это наше действие будет разумно оправданным. *Философия стала анти-априорной*. Но коль скоро мы признали, что большая часть тех истин, которые мы считали априорными, носит контекстуальный и относительный характер, то мы тем самым отвергли единственное веское «доказательство», направленное против теории тождества сознания и тела. Приверженцы теории тождества должны были отметить это обстоятельство, и они не преминули это сделать. Таким образом, ситуация изменилась.

Я использовал понятие «свойство»; однако мне кажется, что существует по крайней мере два понятия «свойство», которые смешиваются друг с другом⁶. Существует весьма древнее понятие, для обозначения которого использовался термин «предикат», как, например, в знаменитом вопросе «Является ли существование предикатом?»; существует также понятие, которое мы используем тогда, когда говорим о «физических свойствах», «главных величинах» и т. д. Когда философ имеет в виду более древнее понятие, он зачастую отождествляет рассуждения по поводу свойств с рассуждениями по поводу *понятий*. Для такого философа *свойства* не могут быть теми же самыми, если утверждение об их тождестве не является концептуальной истиной. Соответственно, свойство «обладать ощущением, имеющим определенный качественный характер», не может быть тождественным свойству «пребывать в определенном состоянии мозга», поскольку соответствующие предикаты не являются *синонимичными* (в широком смысле «аналитической эквивалентности»), а принцип индивидуации для предикатов гласит, что быть *P* — это тот же самый предикат, что быть *Q*, только в том случае, когда «есть *P*» синонимично «есть *Q*».

Рассмотрим, однако же, ситуацию, которая возникает тогда, когда ученый утверждает, что температура — это не что иное, как молекулярная кинетическая энергия. На первый взгляд, перед нами высказывание, утверждающее тождество свойств. Утверждение сводится к тому, что *свойство* обладать определенной температурой является *в действительности* (в каком-то смысле словосочетания «в действительности») *тем же самым свойством*,

6. См.: Putnam H. On Properties // Putnam H. Mathematics, Matter and Method. Philosophical Papers. Vol. 1. Cambridge, 1975.

что и свойство обладать определенной молекулярной энергией; или, выражаясь в более общем ключе, что *физическая величина* температуры — это та же самая физическая величина, что подразумевается молекулярной кинетической энергией. Если это верно, то, поскольку выражение «*x* обладает определенной температурой» не *синонимично* с выражением «*x* нездоровится», означает молекулярную кинетическую энергию», даже когда выражение «нездоровится» является значением молекулярной энергии, соответствующим значению выражения «определенная» для температуры, дело должно обстоять так, что подразумеваемое физиком под выражением «физическая величина» — это что-то совершенно иное, чем то, что философы называют «предикатом» или «понятием».

Точнее, различие в данном случае заключается в том, что в то время как синонимия выражений «*X* есть *P*» и «*X* есть *Q*» требует, чтобы предикаты *P* и *Q* были «тождественными», она вовсе не требует, чтобы были тождественными свойства *P* и *Q*. В противоположность предикатам свойства могут быть «тождественными синтетически».

Если существует *синтетическое тождество свойств*, то почему свойство «обладать нейрофизиологическим состоянием» не могло бы быть *тем же самым свойством*, что свойство «обладать ощущением, имеющим определенное качество» (это очень напоминает ход мысли Спинозы)? Что из того, что данное утверждение не является концептуальной истиной, а многим оно даже кажется ложным *a priori*? В этом заключается суть приведенного аргумента. Короче говоря, мы имеем дело с волной анти-априоризма, мы имеем новый инструментарий синтетического тождества свойств, и с помощью двух этих средств сторонники теории тождества и, в особенности, функционалист, автоматически оказываются в строю.

Теперь же я намереваюсь рассмотреть, что происходит, когда к этим двум средствам прибавляется третье. Что происходит, когда философ является (1) анти-априористски настроенным натуралистом; (2) допускает существование синтетического тождества свойств и (3) и в то же самое время придерживается *последовательно реалистического* взгляда на истину? Я намереваюсь утверждать, что подобный философ столкнется с серьезными эпистемологическими трудностями.

РАСЩЕПЛЕННЫЕ МОЗГИ

Давайте рассмотрим один эксперимент, проведенный нейрофизиологами в последние двадцать лет. Это известный эксперимент, связанный с так называемым «расщеплением мозга». Я хотел бы обсудить ту роль, которую данный эксперимент может иметь для теории тождества, а также представление, которое на протяжении всей дискуссии считалось чем-то само собою разумеющимся, а именно представление о существовании *корреляции*.

Модель мозга как когнитивной системы похожа на компьютер; мозг имеет язык, внутренний язык. Этот язык может быть как врожденным, так и представлять собой своего рода комбинацию, состоящую из врожденного «языка», или системы репрезентаций, и публичного языка. Некоторые философы даже изобрели название для такого гипотетического языка мозга; они называли его «менталез» (*mentales?*)⁷. Давайте рассмотрим, что происходит, когда человек обладает визуальным ощущением в рамках подобной модели. (Признаюсь, я должен выдумать свою собственную неврологию, поскольку я не знаю ее достаточно хорошо, однако я сомневаюсь, чтобы кто-нибудь знал эту науку достаточно хорошо). Вот одна из таких историй:

Когда человек обладает ощущением, выносятся «суждение»; мозг должен напечатать нечто подобное фразе «красное дано в 12 часов». В этом случае качество (назовем его «Q») соответствует, помимо всего прочего, записи на языке «менталез». Кроме того, имеется также импульс, воздействующий на центр, отвечающий за вербальную деятельность, причем этот центр соединен со звукоиздающим устройством, благодаря которому мозг способен произносить на публичном языке: «красное в данный момент». Может получиться так, что суждение, сформулированное на языке «менталез», должно быть перемещено из одного места в другое, прежде чем импульс дойдет до речевого центра. События происходят также и в визуально наблюдаемой коре головного мозга, который исследовался нейрофизиологами

7. Термин введен Джерри Фодором в книге «Язык мысли». См.: Fodor J. The Language of Thought. Harvard University Press, 1975. — *Прим. пер.*

Хьюбелем и Визелем; эти события я представляю себе находящимися на пути к «отчету, сформулированном на языке «менталез»», и к вербальному выражению. Эти «отчеты», «исходные данные» и другие события могут происходить в разных долях головного мозга: если же путем хирургического вмешательства оказывается удаленным *corpus collosum*, правое полушарие головного мозга человека, которое не отвечает за речь, способно увидеть красное (или, по крайней мере будет утвердительно реагировать в ответ на заданный письменно вопрос, опознаваемый только этим полушарием головного мозга). Однако если этого человека спросят, какого цвета карта, то он ответит: «Я не вижу карты». Наконец, имеются определенным образом локализованные мнемонические следы, отвечающие за функционирование памяти. Их можно было бы разделить на *кратковременную, оперативную память* и *долговременную память*. По сути дела речь не идет о наличии какой-то одной линейной каузальной цепи; скорее, речь идет об ответвлениях и соединениях, иными словами, о каузальной сети.

Проблема состоит в том, что психология разделяет ментальные события на отдельные *дискретные* единицы. Вот ощущение голубого. Вот оно появилось; вот оно исчезло. Проблема, однако же, заключается в том, что каузальные сети не являются дискретными. Не существует уникального физического события, которое было бы *конкретным* коррелятом ощущения.

Если верна теория тождества, то состояние ощущения Q тождественно тому или иному состоянию мозга. Метафизический реалист не соглашается с тем, что вопрос, с *каким* состоянием мозга тождественно ощущение Q , можно решить на основе конвенции или свободного выбора. Суть его позиции заключается в том, что *в действительности* мы живем в мире, в котором свойства, с которыми мы имеем дело в опыте как с качественными свойствами ощущений, — это *те же самые свойства*, с которыми мы, правда, иным способом, имеем дело как с физическими свойствами событий, происходящих в мозге. (Или, если выразиться точнее, мы живем в таком мире, в котором свойство обладания ощущением, наделенным определенным качественным характером, в действительности представляет собой не что иное, как свойство находится в определенном нейрофизиологическом состоянии).

Задержимся здесь и рассмотрим, о чем здесь в действительности идет речь. Предположим, что *красное* — субъективное качество, на которое мы обращаем внимание (произведенное, скажем, в результате того, что я смотрел на зеленый диск, а затем этот диск удалили, чтобы я получил остаточный образ). Предположим, что когда я ощущаю это *красное*, то состояние ощущения, в котором я нахожусь, идентично дизъюнкции мозговых состояний. Оно не может быть идентично *одному* максимально определенному мозговому состоянию, потому что мы знаем, что можно убрать любой нейрон, или что-то подобное, и все же иметь тот же опыт. Но это свойство могло бы быть дизъюнктивным — скажем (неправдоподобно), *четно пронумерованные нейроны в затронутой области находятся в возбужденном состоянии или пронумерованные простыми числами нейроны находятся в возбужденном состоянии*. Действительно, это была бы намного более серьезная дизъюнкция. Имелось бы огромное множество неврологических состояний — такое, что их дизъюнкция была бы свойством ощущения *красного цвета*.

Но теперь мы продвинемся немного далее. Если *четно пронумерованные нейроны в затронутой области находятся в возбужденном состоянии*, я испытываю ощущение *красного цвета*. Если цереброскоп говорит: «Нет, это *пронумерованные простыми числами нейроны в затронутой области находятся в возбужденном состоянии*», я все так же испытываю ощущение *красного цвета*. То есть я не могу различить, *в котором из этих мозговых состояний я нахожусь*. Если я испытываю ощущение *красного цвета*, я должен быть в *одном из них*. Но я не могу различать между ними. *Четно пронумерованные нейроны в затронутой области находятся в возбужденном состоянии* — это не наблюдаемое свойство. Даже зная, что теория идентичности является истинной, я не могу вывести из своих ощущений, что я имею это свойство. Назовем это свойство P_1 , а свойство, что *нечетно пронумерованные нейроны находятся в возбужденном состоянии* — P_2 . Состояние ощущения идентично с дизъюнкцией (P_1 или P_2), где это, конечно, третье свойство. P_1 — не состояние ощущения, и P_2 — не состояние ощущения; лишь их *дизъюнкция* — состояние ощущения. Другими словами, в этой онтологии *дизъюнкция двух свойств, которые являются самостоятельно ненаблюдаемыми, может быть наблюдаема*. Именно эту сложную логическую функцию *ненаблюда-*

емых свойств я и ощущаю как простую данность. Такова эта позиция.

Возможно, я выразил этот подход довольно неуклюже. Так, один мой друг заметил: «Предположим, что единственное устройство, которое у нас есть для обнаружения мю-мезонов, не различает мю-мезоны и анти-мю-мезоны. Тогда «быть мю-мезоном» не является наблюдаемым свойством, и «быть анти-мю-мезоном» не является наблюдаемым свойством, но их дизъюнкция является наблюдаемым свойством. Это только кажется парадоксальным для тех, кто считает наблюдаемость менее прагматическим понятием, чем оно есть». Моя цель, однако, состоит не в том, чтобы высмеять подход, который в действительности представляет собой очень важную и законную исследовательскую программу в нейрофизиологии, но в том, чтобы ясно дать понять, к чему он ведет. Я постараюсь показать, что к трудностям ведет не сама по себе теория тождества, но теория тождества в соединении с метафизическим реализмом — то есть в соединении с тем, что я назвал «экстерналистским» взглядом на природу истины.

Можно избежать подобной экстерналистской перспективы. Так, Карнап сказал бы (по крайней мере, на некотором этапе), что разговор о физических объектах выводится за множество ходов из разговора об ощущениях, и что решение сказать, что специфическое мозговое состояние идентично состоянию ощущения *Q*, является в действительности решением изменить язык разговора о физических свойствах таким способом, чтобы изменить наше понятие физического рассматриваемого свойства.

Поскольку речь о физических объектах и физических свойствах выводится за множество ходов из дискурса об ощущениях, мы можем изменять правила. Но это — не точка зрения метафизического реализма, по крайней мере относительно материальных объектов и физических свойств. Тот, кто занимает подобную позицию, может быть метафизическим реалистом в том, что касается *ощущений*, но он не будет метафизическим реалистом относительно *материальных объектов*, и так как он расценивает разговор о материальных объектах как довольно мягкий, он может принимать теорию тождества, просто говоря «я принимаю это как своего рода *соглашение*, как дальнейшее обус-

ловливание значения». Поскольку значения не были полностью установлены заранее, и поскольку сохранялась некоторая открытость структуры, нет никакой проблемы «как мы можем знать, что состояние ощущения тождественно этому свойству, а не с некоторым другим?» Если то, чем *является это* свойство, несколько неопределенно, то мы можем просто постулировать тождество как спецификацию значения. Но я говорю с кем-то, кто действительно думает, что существует материальный мир, и это не выведено из разговора об ощущениях; кто действительно думает, что существуют физические свойства, и кто считает, что такие выражения, как «нейроны в таком-то канале возбуждены», предиктируют наши определенные физические свойства, и эти свойства либо *являются* тождественными этому состоянию ощущения, либо *нет*.

Подобным же образом дело обстоит и для такого философа, как Дэниел Деннет, считающего, что разговор об ощущениях крайне неопределен, и не считающего, что существует хорошо определенное субъективное свойство нахождения в этом состоянии ощущения, свойство наличия ощущения с этим качественным характером. Я думаю, что он также может принять теорию тождества как обуславливание значения, фиксирующую на сей раз не значение терминов физических объектов, а значение психологических терминов. Но опять-таки, это не будет позицией полноценного метафизического реалиста.

Я рассматриваю полноценного реалиста как того, кто думает: «Да, я знаю, каково это психологическое свойство (состояние ощущения). У меня оно есть. Я могу узнавать его. Я думаю, что это — определенное психологическое свойство, на которое я указываю. Я знаю, что такое P_1 и P_2 , а поэтому и то, что такое (P_1 или P_2), и состояние ощущения либо идентично этому, либо нет». Таким же способом наивный физик мог бы сказать: «Нет никакого элемента соглашения» (между прочим, я думаю, что он был бы неправ); «нет никакого элемента соглашения в решении, что *температура* — *средняя молекулярная кинетическая энергия*, поскольку либо температура — это средняя молекулярная кинетическая энергия, либо она — некоторое другое свойство». Это — та точка зрения, которую я хочу исследовать.

Проблема состоит в том, что если мы принимаем эту метафизическую реалистичную точку зрения, то нам открывается

намного больше возможностей, чем люди привыкли считать. Возможность, которая приходит на ум первой, заключается в том, что состояние ощущения идентично свойству наличия соответствующих событий в зрительной коре головного мозга и наличия соответственно зарегистрированной «записи» на языке «менталез», и наличия ввода в речевой центр, и наличия сформированных следов в памяти — то есть о состоянии ощущения можно подумать как об идентичном *соединению* этих нескольких свойств. Но как только мы рассматриваем возможность их разъединения, мы теряем уверенность в том, что нам действительно нужна полная конъюнкция. Возможно, ощущение — это лишь событие в зрительной коре головного мозга? (То есть свойство наличия ощущения — это «в действительности» свойство наличия события в зрительной коре головного мозга.)

Предположим, что это так и есть. Теперь предположим, что мы можем отключить тот процесс, который производит отчет на «менталез», или по крайней мере отключить ввод в речевой центр. Представим себе, что мы показали субъекту красную карту на левой стороне его визуального поля (так, чтобы карта была только «видна только правой доле», как говорят неврологи). Тогда в правой доле зрительной коры головного мозга будет иметь место соответствующее событие, но если мы спросим нашего субъекта: «Вы видите что-нибудь красное?», он ответит «Нет».

Теперь, в соответствии с одним из критериев, которые мы используем, чтобы решить, действительно ли некто имеет ощущение критерием искренних устных сообщений, нам придется сказать, что у *этого субъекта не было ощущения красного*, а тем самым мы *опровергли* теорию о том, что *Q* (релевантная качественная природа) идентична рассматриваемому свойству зрительной коры головного мозга. Но здесь некто может возразить: «Нет, Вы вовсе не опровергли эту теорию. Ведь что это за наблюдатель? У этого парня мозги пополам разделены». Насколько может судить любой наблюдатель *в нормальном состоянии*, *Q* идентично с этим свойством зрительной коры головного мозга. А те наблюдатели, кто не в нормальном состоянии, не считаются. Они *не могут* считаться.

Трудность состоит в том, что есть теории тождества, которые

являются *наблюдательно неразличимыми*⁸, под чем я подразумеваю, что они ведут к тем же самым предсказаниям относительно опыта всех *наблюдателей в нормальном состоянии*.

Рассмотрим подход, согласно которому некто не имеет ощущения красного, если у него нет ввода в речевой центр. Как доказать или опровергнуть это? Можно предположить, что если мы отделим у нашего субъекта ствол головного мозга, но есть некоторая память, которая не проходит устную единицу обработки, тогда у нас есть такой способ. А именно, мы сначала спрашиваем нашего приятеля, имеет ли он ощущение красного. Он говорит, «Нет». Тогда мы приделываем ствол головного мозга обратно (а почему бы и нет!), и спрашиваем: «Ну как, было ощущение красного?» Он мог бы сказать: «Да, но это было какое-то безумие. Вы знаете, что у меня было это ощущение красного, и Вы спросили меня, было ли оно у меня, и я слышал, как я искренне сообщил Вам, что нет». (Мне скажут, что в действительности более общепринято для пациентов «примиряться с судьбой» или рационализировать ситуации этого вида, чем описывать их так, как я только что вообразил.) Разве такое сообщение показало бы, что у субъекта *было* ощущение красного без наличия ввода в речевой центр?

Нет, не показало бы. Если бы Дэниел Деннет (одно время поддерживавший точку зрения, согласно которой ощущение — *это и есть* ввод в речевой центр, или близкий к этому подход⁹) захо-

8. Понятие «Наблюдательной Неразличимости» было введено в статьях по теории пространства-времени Кларком Глаймуром и Дэвидом Маламентом в «Основах Теорий пространства-времени» (*Glymour C., Malament D. Foundations of Space-Time Theories // Minnesota Studies in the Philosophy of Science / Ed. Earman, Glymour, and Srachel. Vol. 8. Minnesota University 1977*). Аналогичная проблема в теории пространства-времени — существование «возможных» пространственно-временных континуумов (то есть континуумов, допустимых в соответствии с теорией относительности), которые отличаются по своим глобальным топологическим свойствам, но в которых наблюдатели имели бы в точности одни и те же опыты. Такие примеры часто отклоняются на том основании, что по «соображениям простоты» мы бы различили, в котором пространственно-временном континууме мы находимся; проблема здесь (как указывает Маламент) состоит в том, что физическая теория (общая теория относительности) не *говорит*, что мы живем в *самом простом* пространственно-временном континууме, совместимом с ее законами.

9. Модель сознания Деннета представлена в статье «К когнитивной теории сознания». См.: *Dennett D. Towards a Cognitive Theory of Consciousness // Dennett D. Brainstorms. Bradford Books, 1978*.

тел согласовать сообщение этого субъекта со своей теорией, то все, что ему для этого потребовалось бы — это сказать «я не отрицаю, что в более позднее время имел место психологическое событие *вспоминания ранее имевшегося ощущения*. Я отрицаю, что *ощущение* имело место в более раннее время». Согласно каждой из этих теорий, субъект *позже* имеет опыт *правильного или неправильного* вспоминания того, что ранее у него было ощущение красного.

Разногласие здесь весьма важно. Большинство неврологов полагает, что у пациентов с «расщепленным мозгом» «сознанием» обладает именно правая доля. В действительности это равносильно тому, чтобы сказать, что иногда есть ощущение красного, или что-то подобное, даже при том, что нет никакого ввода в речевой центр. (Это часто выражается так: «Есть два местоположения сознания»). Однако, по крайней мере один известный невролог, Экклс¹⁰, считает, что отделенная правая доля (или левая доля, в случае тех пациентов, у кого центр речи справа), не обладает сознанием. Согласно подходу Экклса, сознание *унитарно*, то, что отделенная правая доля может «моделировать» сознательное поведение, не показывает, что она — второе местоположение сознания.

Не поможет здесь и обращение к методологическим принципам, например, «выбора более простой теории», поскольку трудно усмотреть уместного вида «простоту», которой обладал бы «унитарный» подход и который отсутствовал бы у подхода «двух местоположений», или наличествовал бы у подхода «двух местоположений» и отсутствовал бы у «унитарного» подхода. Возможно, подход «двух местоположений» более прост в одном отношении; согласно ему, некоторые поведенческие способности (которыми обладает правая доля, даже если она не обладает речью) достаточны для сознания, и это согласуется с фактом, что мы называем животных (которые тоже не обладают речью) сознательными. Но есть много *несходств* между животным с неповрежденным мозгом — чьи мозговые процессы все еще «интегрированы», даже если они не связаны с речью — и частью

10. Сэр Джон Кэрю Экклс (р. 1903) — австралийский нейрофизиолог, с 1968 года работает в США, университет штата Нью-Йорк в Баффало. Нобелевская премия 1963 года. — *Прим. пер.*

«расщепленного мозга». Если бы этот случай не затрагивал нас так непосредственно, если бы у нас не было такой сильной тенденции к метафизическому реализму об *ощущениях*, разве не в полном согласии с нашими лучшими методологическими соображениями было бы расценить это как случай, который *нужно учитывать*, а не бороться с ним?

Короче говоря, есть множество наблюдательно неразличимых теорий идентичности. Если теоретик идентичности прав, то, кажется, в мире нет никакого способа, которым можно знать, каким *способом* он прав; знать, с *каким* мозговым состоянием идентично данное состояние ощущения (или с *каким* мозговым состоянием оно *коррелирует*). Это замечание достаточно важно, чтобы заслужить дальнейшую иллюстрацию.

Томас Нагель¹¹ выдвинул правдоподобное требование, что нельзя представить себе, *каково было бы* быть летучей мышью. Но *почему* это требование должно быть правдоподобно? Несколько лет назад я читал восхитительную книгу о летучих мышах Дональда Гриффина. Я понял, что летучие мыши в основном не отличаются от других млекопитающих. Мы часто думаем, что *можем* представить себе то, какие ощущения имеют наши собаки и кошки. В чем же трудность с летучими мышами?

Летучие мыши могут слышать звуки на несколько октав выше, чем мы. Я не могу представить себе то, *каково было бы* ощущение эхолокации. Но обязательно ли это должно быть *настолько* трудно? Раньше я мог слышать звуки на октаву выше, чем самые высокие звуки, которые я могу слышать в среднем возрасте. Но субъективные высоты тона не изменились: самые высокие звуки, которые я могу слышать, могут быть на октаву ниже, чем самые высокие звуки, которые я мог слышать, когда мне было десять лет, по *объективной* высоте; но самые высокие звуки, которые я слышу, теперь для меня имеют то же самое тонкое, писклявое качество, как и всегда это было со звуками на пороге слышимости. Возможно, именно так летучая мышь и слышит звуки на пять октав выше, чем мы: как короткий высокий писк.

11. Nagel T. What is it Like to Be a Bat? // Readings in Philosophy of Psychology / Ed. Block N. Harvard, 1980.

Теперь вообразим дебаты между двумя философами или психологами, один из которых говорит, что *никакое quale*¹² летучей мыши вообще не подобно никакому человеческому *quale*. *Qualia* летучей мыши невообразимо отличаются от человеческих *qualia*. Вы никогда не сможете представить себе, что испытывает летучая мышь (или даже собака или кошка). Другой философ, как мы можем вообразить, отвечает: «Ерунда! Возможно, есть *некоторые* ощущения летучей мыши, которые я не могу представить себе. Некоторые ощущения других людей (например, некоторые ощущения людей другого пола) я, вероятно, тоже не могу себе представить, но это не подразумевает, что я расцениваю психологическое пространство этих других людей как невообразимо отличающееся от моего собственного. Почему же я не должен думать, например, о визуальной области летучей мыши как об очень похожей на мою собственную визуальную область? (N. B. Вопреки легендам, летучие мыши очень хорошо видят). С учетом некоторых поправок на оптику глаза летучей мыши или слуха летучей мыши в пределах диапазона, который накладывается на мой, ее слух подобен моему, а ее боль подобна моей боли». Можем ли мы *уладить* этот спор?

Поскольку у человека и летучей мыши различно число нейронов и их организация (акустический центр мозга летучей мыши достигает 7/8-ых его размера), свойства на наиболее определенном неврологическом уровне — число возбужденных нейронов, — идентичные с *quale* в случае летучей мыши (при условии, что теория идентичности правильна), не могут быть буквально теми же самыми свойствами, которые идентичны с каким бы то ни было *quale* в случае человека. Или же могут? Предположим, что когда летучая мышь имеет некоторое визуальное ощущение (при виде красных объектов), то ее мозг имеет дизъюнктивное свойство (P_1 или P_2), где P_1 и P_2 являются максимально определенными состояниями мозга летучей мыши. (Это было бы в действительности гораздо более сложное дизъюнктивное свойство с тысячами случаев, но давайте упрощать.) И предположим, что, когда у меня есть некоторое визуальное ощущение

12. Свойство, качество (лат.). Патнэм употребляет латинский термин, чтобы подчеркнуть особый характер обсуждаемых здесь свойств — это качественные характеристики сферы ментального. — Прим. пер.

(при виде красных объектов), мой мозг имеет дизъюнктивное свойство (P_1' или P_2'). Рассмотрим следующие две теории: (1), что качественный характер ощущения летучей мыши (назовем его «*красное_{ЛМ}*») идентичен с (или *по крайней мере коррелирует с*) дизъюнктивным свойством (P_1 или P_2), а качественный характер человеческого ощущения (назовем его «*красное_ч*») идентичен с (или *по крайней мере коррелирует с*) другим свойством (P_1' или P_2'); (2), что качественный характер ощущения летучей мыши *идентичен* с качественным характером моего ощущения (то есть *красное_{ЛМ} = красное_ч*) и оба *идентичны* с (или *по крайней мере коррелируют с*) более сложным дизъюнктивным свойством (P_1 или P_2 или P_1' или P_2').

Согласно первой теории, летучая мышь и я имеем различные опыты, в то время как согласно второй у нас с ней одинаковый опыт; но эти две теории ведут к одним и тем же предсказаниям относительно того, что испытают *человеческие* наблюдатели, нормальные и ненормальные. Еще раз повторим: они наблюдательно неразличимы.

Помогут ли здесь методологические принципы («выбор более простой теории»)? И вновь не ясно, помогут ли. Нед Блок заметил, что первая теория более проста в одном отношении (*qualia* идентифицировано с более простым физическим свойством в *каждом* случае), но зато вторая более проста в другом отношении (вторая теория — «не шовинистская»; она допускает, что не обязательно иметь *в точности* нашу физическую конституцию, чтобы иметь наши *qualia*). И снова нам не хватает принципов определения уникального привилегированного обмена. Действительно, чей разум должен решать о существовании и статусе таких принципов? Почему бы нам, как настаивал Витгенштейн, не отказаться от нашего метафизического реализма по вопросу об ощущениях и о «том же самом» (в применении к ощущениям), и не обращаться также и с этим как со случаем, который нужно учитывать, а не бороться с ним?

Наконец, я хочу представить три теории, которые, как я уверен, являются ложными, но которые трудно или невозможно исключить, если метафизический реализм прав. Они таковы: (1), что *красное_ч* в конце концов идентично с функциональным (или квазифункциональным) состоянием, а именно с состоянием нахождения в любом материальном (например, физическом)

состоянии, которое раньше всего в вашей жизни играло функциональную роль нормативной передачи сигналов присутствия объективно красного; (2), что камни имеют *qualia* (то есть в камнях имеют место события, качественно подобные, насколько это возможно, *визуальным ощущениям*); (3), что нации обладают сознанием.

Сначала рассмотрим (1). Вспомним аргумент, с помощью которого я показал, что *красное_c* не могло бы быть функциональным состоянием. Этот аргумент состоял в том, что если мы идентифицируем *красное_c* с функциональным состоянием *нахождения в любом материальном состоянии (например, определенном состоянии мозга), которое обычно сигнализирует о присутствии объективно красного*, то я не подвергся бы инверсии спектра (по крайней мере в случае «амнезии»), так как я нахожусь в этом функциональном состоянии, когда вижу нечто объективно красное, и *до*, и *после* инверсии спектра (отводя время для лингвистического регулирования, и, если необходимо, постулируя приступ амнезии). Но согласно метафизической реалистической позиции, конечно, возможно, что я *подвергся* инверсии спектра (даже при том, что я не помню об этом из-за амнезии). Случай даже более убедителен, если я не подвергся амнезии и *вспоминаю*, что мой спектр был перевернут; даже в этом случае, если лингвистическое урегулирование стало автоматическим, есть смысл, в котором то, что было ранее «ощущением зеленого», теперь играет функциональную роль «передачи сигналов присутствия в окружающей среде объективно *красного*».

Этот аргумент показывает лишь, что *красное_c* не идентично с упомянутым функциональным состоянием. Он не показывает, что это не идентично с более сложным функциональным состоянием, вроде *состояния нахождения в любом материальном состоянии, которое раньше всего в вашей жизни реализовало вышеупомянутое функциональное состояние*. Можно возразить, что это — забавное свойство, сложная логическая функция функциональных свойств. Но почему сложная логическая функция функциональных свойств с меньшей вероятностью будет идентичной с *красное_c*, чем дизъюнкция сложных физических свойств? Предпочитает ли мир дизъюнкции физических свойств конъюнкциям функциональных свойств?

Рассмотрим (2). Пусть P_c — свойство быть камнем, и рассмот-

рим гипотезу, что *красное*, идентично с дизъюнктивным свойством (P_1 или P_2 или P_1' или P_2' или P_3). Конечно, камни постоянно имеют это свойство. Так, согласно этой гипотезе, события качественного характера *красное*, имеют место в камнях постоянно. (Они не испытывают ощущение красного в функциональном смысле испытания ощущения красного, но в них постоянно имеет место событие того же качественного характера, что и то событие, которое играет в нас функциональную роль быть ощущением красного.) Или же рассмотрим более сложные гипотезы, согласно которым камни имеют различные *qualia* в разное время. Или же только гипотезу о том, что какая-то этих гипотез (не указывая, которая) правильна. Мы могли бы сказать: «Да эти гипотезы сумасшедшие!» Да, они таковы. Но каждая из них ведет к тем же самым предсказаниям относительно всех человеческих наблюдателей, что и «нормальная» теория. Ни одна из них не может быть исключена на наблюдательных или экспериментальных основаниях, потому что каждая из них наблюдательно неотличима от более стандартного подхода.

Мы могли бы подумать, что эти теории могут быть исключены обращением к методологическому принципу, согласно которому *нельзя приписать свойство объекту, лишенному разума*. Конечно, этот принцип не утверждает, что эти теории *ложны* (иногда вещи, которым мы не имеем никаких причин верить, истинны), но, по крайней мере, он говорит, что у нас есть *обоснование* для того, чтобы считать их ложными. Но *действительно* ли нет никакой причины поддержать наименее определенную из этих теорий (теорию о том, что некоторая такая теория правильна, и камни имеют *qualia*)? Если мы имеем *qualia*, и физикализм истинен (а многие философы думают, что есть много серьезных оснований для принятия физикализма), то есть по крайней мере один физический объект, в котором имеют место события с качественным характером: тогда почему такие события не должны иметь место во *всех* физических объектах? Если мы могли бы показать, что непосредственно относительно *quale* есть нечто такое, что *требует*, чтобы оно имело специфическую функциональную «роль», которую оно имеет лишь у людей, то этот ход был бы заблокирован; но именно об этом те, кто верит, что *qualia* — это метафизически реальные объекты, сообщают нам, что мы не можем этого сделать.

И, наконец, рассмотрим (3). Рассмотрим гипотезу, что боль является идентичной с соответствующим функциональным состоянием, которое может быть показано или организмами, или нациями. Другими словами, предположим, что, когда Соединенные Штаты объявляют, что «Соединенные Штаты испытывают боль...», то именно это действительно и происходит. Мы, конечно, никогда не догадались бы об этом. Возможно, читатель в этот момент находит интересным и умеренно забавным предположение о том, что группа может вести себя такими способами, которые напоминают те способы, которыми ведет себя нечто, что действительно чувствует боль, когда это нечто проявляет свою боль; но читатель не думает, что Соединенные Штаты действительно чувствуют боль. Согласно этой гипотезе, читатель был бы неправ: национальный *Geist*¹³ действительно чувствовал бы боль.

Эта гипотеза связана с интересным обсуждением в философии сознания. Аргумент, который любят использовать функционалисты (включая меня) — это «антишовинистический» аргумент: в принципе, различия между роботом и человеком (во всяком случае, в их функциональной организации) могли бы быть сведены до мелких деталей физики и химии. Возможен даже робот, соответствующий нам вплоть до нейронного уровня. (Он мог бы даже иметь тело из «плоти и крови», кроме мозга.) Различие состояло бы в том, что наши нейроны сделаны из углерода, водорода, белков и так далее, а у робота они бы были электронными, но вся схема, начиная с нейронного уровня, точно соответствовала бы нашей. Теперь, если вы не «водородно-углеродистый шовинист», думающий, что углероду и водороду более свойственно сознание, то почему бы вам не считать, что этот робот — человек, чей мозг просто содержит больше металла и меньше водорода и углерода?

Этот аргумент вызвал следующий ответ: «Хорошо, но давайте вместо этих электронных устройств — электронные нейроны, соединенных в те же самые схемы, что и человеческие нейроны — предположим, что у нас есть такие миниатюрные человечки, маленькие герлскауты и бойскауты». Нам даже не надо представлять, что эти маленькие человечки вообще знают то, для чего предназначена схема в целом, или что они видят что-

13. Дух (нем.). — Прим. пер.

либо, кроме смутно освещенной комнаты, или многих смутно освещенных комнат, в которых они передают сигналы друг другу. (Конечно, их время должно было бы идти гораздо быстрее «нашего» времени.) Они могли бы быть отчужденными работниками. «Теперь, — продолжается ответ, — вы не назвали бы эту вещь «сознательной», потому что вы знаете, что на самом деле это только маленькие человечки, перемещающие тело. И это показывает, что соответствующей функциональной организации (подобной нашей) недостаточно для того, чтобы обосновать применение таких предикатов как «сознательный»».

Возможное возражение на этот ответ (в действительности сделанное мной) отрицает, что «гидроголовый робот» (как я назвал эту вещь) *имеет* ту же самую функциональную организацию, что и мы. Но я мог бы дать и более радикальный ответ. Я мог бы сказать: «Почему мы *не должны* называть гидроголового робота сознательным? Если первый аргумент правилен (а я думаю, что это так), если бы робот с позитронным мозгом обладал сознанием, то почему тот факт, что нейроны гидроголового робота в *большей степени* обладают сознанием, должен означать, что система в целом обладает сознанием в *меньшей степени*? В конце концов, каждый из нас — в некотором смысле общество маленьких животных, так как наши клетки — в некотором смысле индивидуальные животные. И, возможно, у них есть какие-то чувства, кто знает? Помимо *наших* чувств». Теперь, если мы пойдем по этому пути, если мы решим, что гидроголовый робот обладает сознанием (даже при том, что его нейроны — бойскауты и герлскауты), тогда почему бы Соединенным Штатам не обладать сознанием?

Конечно, я не утверждаю, что Соединенные Штаты имеют *ту же самую* функциональную организацию, что и *homo sapiens*¹⁴. Ясно, что это не так. Но есть много подобий. У Соединенных Штатов есть защитные органы. У них есть органы глотания, они потребляют нефть, медь и так далее. Они производят обширные выделения, загрязняя окружающую среду. Не подобно ли это функциональной организации млекопитающего по крайней мере *настолько же*, насколько у извивающейся мухи, которой мы *приписываем* боль?

14. Человек разумный (лат.). — Прим. пер.

НАСКОЛЬКО ХОРОШО ОПРЕДЕЛЕН «КАЧЕСТВЕННЫЙ ХАРАКТЕР»?

Что означает сказать, что два ощущения одного индивида имеют или не имеют один и тот же качественный характер? До сих пор мы не оспаривали, что это совершенно ясно. Но это не так даже на интроспективном уровне. Ведь, с одной стороны, именно то, чем индивиду кажутся его переживания, печально известно своей зависимостью от предшествующей концептуализации, как в случае, когда мы сообщаем, что видим круглую крышку стола, несмотря на то, что смотрим на нее сбоку.

Применительно к случаю с круглой крышкой стола, начиная с XIX столетия, психологи и философы обсуждали две конкурирующих теории: либо индивид имеет «эллиптические чувственные данные» и *думает*, что они круглые (если только это не «натренированный интроспекционист»), либо он имеет круглые «гештальты» и только *думает*, что они эллиптические, вследствие оптической теории. Возможен индивидуальный опыт, удовлетворяющий каждому из этих описаний; и многие опыты будут удовлетворять любому из них. Вряд ли неврология способна разрешить этот спор: эллиптический образ на сетчатке, несомненно, является причиной событий в самом мозгу, и если мы отождествляем *их* с «визуальным ощущением», то в результате мы вполне можем получить что-то вроде классической истории про «эллиптические чувственные данные плюс бессознательные выводы». Обсуждаемый характер опыта («Я вижу круглую крышку стола») также имеет соответствия среди «записей» и «входных данных» мозга, и если мы отождествляем *их* с визуальным ощущением, мы вполне можем получить историю, в которой индивид не имеет эллиптических чувственных данных, если только он не *выносит суждение*, что нечто выглядит эллиптически. Почему бы нам не сказать, что обе эти версии законны, как говорит Гудмен, комментируя случай с субъектом, которого попросили описать видимое движение,

Лучшее, что мы можем сделать, это определить вид терминов, словарь, который он должен использовать, говоря ему, чтобы он описывал то, что он видит, в терминах восприятия или феноменов, а не в физических терминах. Независимо от того, вызывает это или нет различные реакции, это проливает совер-

шенно новый свет на то, что происходит. Тот факт, что инструмент, который предполагается использовать для моделирования фактов, должен быть определен, делает бессмысленным любое отождествление физического с реальным, а перцептивного — со всего лишь видимым. Перцептивное — не в большей степени весьма искаженная версия физических фактов, чем физическое — крайне искусственная версия перцептивных фактов¹⁵.

Если я вижу красную скатерть дважды на протяжении одного дня, имею ли я *одно и то же* ощущение красного? Или я имею разные ощущения и не замечаю различия (если только, случайно, я не художник)?

Особенно запутывающим существо дела случаем является случай *аккомодации*. Если субъекту даны очки со стеклами, переворачивающими образ, то через какое-то время вещи опять будут для него выглядеть нормальными. Это чувственные данные «встали на место»? Или субъект приспособился к измененным чувственным данным путем переинтерпретации понятий «верх» и «низ»? Весьма вероятно, что сам субъект не может сказать, когда именно все пришло в норму или что именно из перечисленного случилось. (Те из читателей, кто, как и я, носят двухфокусные очки, могут спросить себя: не выглядит ли нижняя часть визуального поля по-другому, даже когда различие остается незамеченным?) Поскольку существуют изменения, к которым субъектам никогда не приспособится (фактически, аккомодация происходит только в случаях относительно простых изменений), а я утверждал, что субъект не приспособится к инверсии цвета, феномен аккомодации, разумеется, порождает сомнения в том, насколько хорошо определено понятие «тот же качественный характер».

РЕАЛИЗМ В ВОПРОСЕ О QUALIA

Мы рассмотрели набор скептических замечаний. Они нацелены продемонстрировать не ложность теории тождества (или теории корреляции — заметим, что все эти трудности могут в

15. Goodman N. The Ways of Worldmaking. Hackett Publishing Company, 1978. PP. 92–3. [Ср.: Гудмен Н. Способы создания миров / Пер. с англ. М. В. Лебедева. М.: Идея-Пресс, Праксис, 2001. С. 207. — Прим. пер.].

равной степени утверждаться и как трудности «корреляционистского» взгляда), но что, если она истинна, то существует настолько большое число альтернативных способов определения деталей, что никогда нельзя будет знать, какой же из них истинный. А незнание последнего означает незнание ответа на огромное число традиционных скептических вопросов, таких, как: имеют ли *qualia* камни и другие неодушевленные объекты, имеют летучие мыши и другие виды животных такие же *qualia*, что и мы, или нет, могут ли группы чувствовать боль и так далее.

Но почему какой-либо философ должен полагать хотя бы логически возможным, что камень может испытывать боль (т. е. что событие того же «качественного характера», что и человеческая боль, может случиться «в» камне)? Возможно, Рассел дал нам некий ключ к природе этого вида метафизического реализма. Рассел был реалистом в вопросе о *qualia* и реалистом в вопросе об универсалиях. Более того, он считал *qualia* образцовыми универсалиями. Универсалия есть, прежде всего, способ, каким вещи могут быть *подобными*; и Расселу казалось, что качественные подобию индивидуальных ощущений представляют собой эпистемологически наиболее примитивные и наиболее фундаментальные примеры «способов, какими вещи могут быть подобными». *Qualia* для Рассела суть универсалии по преимуществу.

Традиционный реалист, однако, полагает универсалии полностью хорошо определенными: *слова* могут быть смутными, но сами универсалии не могут быть смутными. (Как сказал однажды Гёдель в разговоре, смутное слово таково, потому что оно замещает смутный набор понятий, но понятия совершенно отчетливы.)

Поэтому, если *qualia* суть универсалии, а универсалии по своей природе отчетливы, должно быть совершенно отчетливо, проявляет любая данная вещь или любое данное событие — включая половину расщепленного мозга или некое событие в нем, камень или некое событие в нем, нацию или группу или некое событие в ней — или нет любое данное *quale*. И если это *quale* мыслится как независимое от функциональной роли, которую оно играет, если мыслится полностью случайным, что качественный характер ощущения красного есть качественный характер чего-то, что выполняет эту конкретную функциональ-

ную роль, то тогда, кажется, должна быть логическая возможность того, что расщепленный мозг или камень имеют это *qualé*.

Философ, наподобие меня, желающий отрицать, что каждая из этих возможностей имеет смысл (хотя некоторые из них могут — *существует* искушение считать правое полушарие расщепленного мозга «местоположением сознания» и я утверждал, что было бы законно решить так поступить), должен убедиться, что он ясно дал понять, что не исповедует какой-либо формы бихевиоризма. Сказать, что *qualia* не являются хорошо определенными сущностями — это не то же самое, что сказать, будто они не существуют, что это все только поведение или что бы то ни было еще. Многие понятия смутны и тем не менее имеют некоторых ясных референтов. Например, понятие *дома* не является четким относительно случая иглу (иглу — это дом или не дом?), случая хоганов, возможно также в других случаях. Но тот факт, что никакой факт в мире не может позволить нам решить, является ли иглу домом, не означает, что домов не существует. И, сходным образом, тот факт, что никакой факт в мире не может позволить нам решить, «сознательно» ли правое полушарие, не означает, что не существует сознательных существ.

Качественное подобие ощущений до некоторой степени определено: если я имею ощущение красного, за которым следует ощущение зеленого, я знаю, что я имел несходные ощущения (и я знаю это без сравнения их функциональных ролей), а если я имею ощущение красного, за которым следует «такое же» ощущение красного, я знаю (в зависимости от смутности, о которой говорилось выше), что имел подобные ощущения. Но для того, кто смотрит на истинность из «интерналистской» перспективы, из сказанного не следует, что в каждом случае есть какое-то положение дел, которое указывало бы нам, являются ли два ощущения качественно подобными или неподобными (не говоря уже о двух произвольно взятых событиях).

Пусть *E* — событие обладания мною определенным ощущением в определенное время, а *E'* — некое физическое событие, имеющее место в камне. Предложение, что качественный характер *E* (скажем, *красное*„) мог бы быть тождественен или коррелятивен некоему свойству такого вида: (*P*₁ или *P*₂ или *P*₃) (где *P*₁ — свойство *бытия камнем*) — оскорбляет всякое человеческое здравомыслие. Предложение, что *E* и *E'* могли бы быть «качествен-

по подобным» событиями, абсурдно. Мы уже обсуждали одно объяснение этой абсурдности: объяснение, согласно которому эта гипотеза абсурдна, поскольку она нарушает методологическую максиму «не приписывай без основания свойства объекту». Даже если бы это объяснение работало, из него следовало бы много меньше, чем невозможность для камней иметь *qualia* (или некогерентность идеи, что они их имеют). Если только *это* не в порядке с «гипотезой», что камни имеют *qualia*, то мы можем быть обязаны утверждать следующее: *согласно всему, что нам известно, возможно, что камни имеют qualia, но в высшей степени невероятно a priori, что они их имеют.*

В действительности, гипотеза, что камни имеют *qualia*, не когерентна в значительной степени в том же смысле, в каком не когерентна гипотеза Мозга в Бочке: она также предполагает магическую теорию референции. Любое вменяемое человеческое существо считает *Е* и *Е'* настолько неподобными, что не возникает даже вопроса о «качественном подобии» (в том смысле, в каком два ощущения могут быть качественно подобными, т. е. одинаково ощущаться). Но метафизический реалист, хотя и не отрицая этого полностью, полагает, что *Е* и *Е'* могли бы быть так подобны (что это логически возможно), *несмотря даже на то, что это «сумасшедший» образ мысли.* И он полагает это, поскольку находится во власти иллюзии, что обладание таким ощущением с его качественным характером, «тем, как оно ощущается», с его функциональной ролью, с сопутствующими ему мыслями и суждениями, каким-то образом является причиной того, что выражение «то, как это ощущение ощущается» (или его технический заменитель, например: «качественный характер этого ощущения» или «красное», или «эта *quale*») указывает на одну определенную «универсалию», одно абсолютно отчетливое свойство метафизических индивидуальных событий. Но это не так.

Если бы на самом деле существовали роботы, функционально изоморфные нам, и мы бы с ними работали, спорили, дружили бы с некоторыми из них, мы бы быстро почувствовали уверенность в том, что они обладают сознанием. (Мы даже могли бы быть озадачены вопросом, а не имеют ли они те же *qualia*, что и мы; но мы не думали бы об этом чаще, чем мы раздумываем над вопросом: не имеют ли летучие мыши или собаки *такие*

же *qualia*. что и мы?) Предположим, однако, что нам повстречались гидроголовые роботы. (Вообразим, что они действительно *развились* таким образом благодаря какому-то биологическому процессу, имевшему место где-то, точно так же, как развиваются на Земле животные, находящиеся в симбиотических отношениях.) Что бы мы подумали о *них*?

Хотя на самом деле в столь странном случае ни в чем нельзя быть уверенным, похоже, что даже в таких обстоятельствах (если бы мы в основном взаимодействовали с целым роботом и только изредка с его сознательными «нейронами» — «бойскаутами и герлскаутами» из моей истории) мы могли бы начать приписывать сознательность. Но, вероятно, мы навсегда остались бы разделенными на две противоположные точки зрения. Если бы мы, в конце концов, обрели уверенность в том, что гидроголовые роботы сознательны, то могли бы мы тогда начать быть крайне мало щепетильными в отношении Соединенных Штатов? Я не знаю.

Перспектива, которую я настоятельно рекомендую, учитывая все эти случаи, заключается в том, что здесь нет ничего *спрятанного*, никакого поуменального факта о действительной сознательности сущностей или их действительной несознательности, или действительном подобии качеств или различии между ними. Есть только очевидные эмпирические факты: что камни и нации сильно отличаются от людей и животных; что роботы разных видов представляют собой переходный род объектов между людьми и животными, с одной стороны, и камнями и нациями, с другой; и так далее. Камни и нации *не* имеют сознания; это — факт, касающийся того понятия сознания, которым мы реально располагаем.

Такая линия рассуждений кажется столь возмутительной в силу того, что она ставит наши стандарты рациональной приемлемости, обоснования и, в конечном счете, истины в зависимость от стандартов подобия, которые, очевидно, представляют собой продукт нашего биологического и культурного развития (например, некоего факта, что взаимодействовали или не взаимодействовали с «интеллектуальными роботами»). Но что-то в этом роде истинно относительно большей части языка, используемого нами в повседневной жизни, относительно таких слов, как «личность», «дом», «снег» и «коричневый», например.

Реалист, принявший это решение головоломок *qualia*, вероятно, выразил бы его, сказав, что «*qualia* не существуют в действительности» или что *qualia* принадлежат к нашей «концептуальной схеме второго класса»; но какой смысл может иметь такое понятие «существования», которое относит *дома* к *несуществующему*? Наш мир — человеческий мир, и что является, а что не является сознательным, что имеет ощущения, а что их не имеет, что чему качественно подобно, а что чему нет — все это, в конечном счете, зависит от наших человеческих суждений о сходстве и различии.

5. ДВЕ КОНЦЕПЦИИ РАЦИОНАЛЬНОСТИ

В предыдущей главе речь шла о рациональности и «рациональной приемлемости». Однако рациональность не так-то просто поддается объяснению.

Эта проблема имеет определенные параллели и в иных областях. Несколько лет тому назад я исследовал поведение слов, используемых для обозначения естественных видов, например, слова «золото», и пришел к заключению, что объем термина не определяется простым «набором семантических правил» или какими-либо иными институциональными нормами. Нормы могут определять, что определенные объекты являются *парадигматическими примерами* золота; но они не определяют полный объем термина. Более того, нет ничего невозможного в том, что даже парадигматический пример мог бы оказаться чем-то, что не является золотом, как это произошло бы в том случае, если бы нормы просто *определяли*, что значит быть золотом.

Мы готовы считать какой-нибудь предмет принадлежащим к определенному виду (даже если наших нынешних проверок недостаточно для того, чтобы показать, что этот предмет принадлежит к этому виду), тогда, когда этот предмет имеет по существу ту же самую природу (или, более смутно, является «в достаточной степени сходным»), что и парадигматические примеры (или большинство из них). Что собой представляет существенная природа и что считать достаточным сходством, зависит как от естественного вида, так и от контекста (чай со льдом может быть «водой» в одном контексте, но не в другом); однако в случае с золотом во внимание принимается именно его элементарный состав, поскольку со времен древних греков именно ему отводилась роль фактора, определяющего законосообразное поведение субстанции. Если мы не согласимся с тем, что то вещество, которое греки принимали за *chrysos*, чем бы оно ни было, имело ту же самую сущностную природу, что и парадигматические образцы, то для нас будут лишены всякого смысла как поиск греками новых методов для обнаружения подделок под золото (который привел Архимеда к изобретению проце-

дуры проверки удельного веса золота), так и их физические размышления.

Возникает соблазн поступить точно так же и с проблемой рациональности и показать, что фактором, определяющим, является ли убеждение рациональным, являются не нормы рациональности той или иной культуры, но *идеальная теория рациональности*, теория, которая обеспечивает необходимые и достаточные условия рациональности убеждения при соответствующих обстоятельствах в любом возможном мире. Такая теория должна была бы *объяснить* парадигматические примеры, точно так же как идеальная теория золота объясняет парадигматические примеры золота. Однако она могла бы пойти и дальше и предоставить нам критерии, которые позволили бы понять случаи, до сути которых мы в данное время не состоянии добраться, подобно тому как наша современная теория золота позволяет нам понимать случаи, которые поставили бы в тупик даже самого проницательного древнего грека. Самое важное затруднение, возникающее при попытке толковать термины «рациональный», «разумный», «обоснованный» и т. д. как термины естественных видов, состоит в том, что мы не располагаем сколько-нибудь обнадеживающими перспективами *найти* заслуживающие доверия обобщения для всех рационально приемлемых убеждений. Все отдельные частицы золота подчиняются эффективным универсальным законам, и именно это обстоятельство позволяет описывать золото как вещество, поведение которого будет подчиняться этим известным нам законам. Однако каковы наши шансы найти эффективные универсальные обобщения, которым бы подчинялись все рационально обоснованные убеждения?

То обстоятельство, что эти шансы являются не слишком уж обнадеживающими, вовсе не означает, что не существует *никаких* аналогий между научным исследованием природы логики и моральным или философским исследованием. Например, в этике мы начинаем с суждений о том, что отдельные действия являются правильными или неправильными (эти суждения играют, так сказать, роль «отчетов о наблюдениях»), а затем на основе этих суждений постепенно формулируем *принципы поведения* (заметим, что они не являются такими обобщениями, которые не допускают исключения). Эти принципы поведения часто сопро-

воздаются какими-то соображениями или же примерами, призванными играть роль иллюстрации, как, например, в случае с суждением: «Не обижайте странников, поскольку Вы знаете, каково быть странником в Египте», которое представляет собой обобщение высокой степени общности. В свою очередь, эти максимы, эти принципы поведения воздействуют на нас и изменяют наши суждения об отдельных случаях, результатом чего может быть появление новых максим, дополняющих или модифицирующих старые. По прошествии тысячелетий этой диалектики принципов поведения и суждений об отдельных случаях, может появиться философ и предложить моральную концепцию («теорию»), которая может изменить как принципы поведения, так и суждения об отдельных случаях и т. д.

Та же самая процедура может быть обнаружена и в философии как таковой, которая по сути дела совпадает с теорией рациональности. Несколько лет назад в одной из своих публикаций¹ я, следуя Грайсу и Беккеру, описал недостающие моральной теории элементы и включил в их состав (1) пожелание, чтобы основополагающие допущения обладали по меньшей мере *широкой привлекательностью*; (2) пожелание, чтобы система была способна выдержать рациональную критику и (3) пожелание, чтобы моральные рекомендации были *жизненными*.

Сходным образом, добиться лучшего понимания природы рациональности и притом единственным способом, нам известным, — можно лишь развить лучшие философские концепции рациональности. *Подобного рода развитие представляет собой бесконечный процесс, однако иным оно и не могло бы быть.* Поразительно, но те недостающие моральной системе элементы, что были перечислены мною выше, с равным правом можно было бы привести в качестве элементов, недостающих методологии или системе рациональной процедуры в любой области человеческих интересов. В аналитической философии основные попытки добиться лучшего понимания природы рациональности подобного рода способом предпринимались в рамках философии науки, результатом чего стало появление двух важных тенденций.

1. Putnam H. Literature, Science and Reflection // New Literary History. vol. 7. 1975–1976. Перепечатано в сборнике: Putnam H. Meaning and Moral Sciences. L.: Routledge and Kegan Paul, 1978.

ЛОГИЧЕСКИЙ ПОЗИТИВИЗМ

В последние пятьдесят лет наиболее ярким проявлением тенденции считать методы «рационального оправдания» своего рода заданным набором стандартов или критериев (пусть даже пока еще не получившим адекватной формализации со стороны философов науки) было движение, известное как «логический позитивизм». Причем логические позитивисты рассчитывали не только на то, что в один прекрасный день «логикам науки» (так позитивисты называли философов) удастся сформулировать список или критерий, который, как предполагалось, исчерпывающим образом описывал бы «научный метод»; они полагали также, что коль скоро «научный метод» и рациональность — это по сути дела одно и то же и коль скоро проверяемость при помощи научного метода определяет осмысленность («Значением предложения является метод его верификации»), то список или критерий определял бы, какие высказывания являются познавательно значимыми, а какие — нет. Высказывания, поддающиеся проверке при помощи методов, внесенных в список, то есть при помощи методов математики, логики и эмпирических наук, считались бы в этом случае осмысленными; все же прочие высказывания, как полагали позитивисты, являются «псевдо-высказываниями» или замаскированной бессмыслицей.

Очевидное возражение заключалось бы в утверждении, что критерий значимости логического позитивизма является *самоопровергающимся*: поскольку этот критерий сам по себе не является ни (а) «аналитическим» (термин, который позитивисты использовали для объяснения логики и математики), ни (б) эмпирически проверяемым. Как ни странно, эта критика не оказала существенного влияния на логических позитивистов и в очень незначительной степени притормозила распространение этого движения. Я думаю, что пренебрежение этим философским гамбитом было большой ошибкой и что этот гамбит является не только правильным, но и поучительным, причем не только в случае логического позитивизма.

Та мысль, которую я собираюсь развить, будет опираться на следующее наблюдение: формы «верификации», допускаемые логическими позитивистами, — это формы, которые были *ин-*

ституционализированы современным обществом. То, что поддается «верификации» в позитивистском смысле этого слова, может быть верифицировано, поскольку, в зависимости от обстоятельств, является или правильным (в нефилософском или в дофилософском смысле этого слова), или вероятно правильным, или же успешным в науке. Общественное же признание правильности, или вероятной правильности, или же статуса «большой степени успешности научной теории» служит иллюстрацией тех образов знания и норм рациональности, которые поддерживаются нашей культурой, что пропагандирует и укрепляет их.

Первоначальная позитивистская парадигма верификации не была общественно институционализированной парадигмой. В работе Карнапа «*Der Logische Aufbau der Welt*» («Логическое построение мира») верификация считалась приватной, основанной на ощущениях, субъективное качество или «содержание» которых полагалось «недоступным для сообщения другому лицу». Однако под влиянием Нейрата Карнап вскоре перешел к более общественной, более «интерсубъективной» концепции верификации.

Поппер акцентировал идею, согласно которой научные предсказания сопоставляются с «базисными предложениями», предложениями типа «правая чашка весов пошла вниз», принимаемыми на основе общего согласия, даже если они, к удовлетворению скептиков, не могут быть доказаны. В связи с этим его критиковали за использование конвенционалистского языка, поскольку у него получалось, что базисное предложение принимается на основе конвенции или социального решения. Однако по моему мнению, то, что в учении Поппера кажется конвенционалистским элементом, в действительности было просто признанием *институционализированной* природы неявных норм, к которым мы обращаемся, вынося обыденные суждения восприятия. Наш ответ скептику, требующему от нас доказательства такого высказывания, как «я стою на полу», свидетельствует о существовании социальных норм, требующих согласия с подобного рода высказываниями при соответствующих обстоятельствах.

Витгенштейн утверждал, что без таких общественных норм, норм, разделяемых группой и составляющих «форму жизни», язык и даже мышление были бы невозможными. Согласно Вит-

генштейну, абсурдно спрашивать, является ли институционализируемая верификация, о которой я вел речь, «реально» обосновывающей. В своей работе «О достоверности» Витгенштейн отмечает, что философы могут дать сто одно эпистемологическое «обоснование» высказывания «кошки не растут на деревьях», но ни один из них не начнет свое рассуждение с более надежной (в институционализированном смысле слова «надежный») отправной точки, чем та, которую даст нам факт, что кошки не растут на деревьях.

Скептики выражали сомнение не только по поводу суждений восприятия, но и по поводу обыкновенной индукции. Юм, у которого я не нахожу разницы между тем, что *рационально*, и тем, что *разумно*, указывал, что не существует *рационального* доказательства того, что в Соединенных Штатах этой зимой будет идти снег (или, хотя бы, что вероятно будет идти снег), хотя он при этом бы и добавил, что было бы весьма неразумно сомневаться в том, что снег пойдет. Однако наш ответ скептику, заставляющему нас «доказать», что этой зимой в Соединенных Штатах пойдет снег, состоит в указании на общественные нормы, требующие согласия по подобного рода «индуктивным выводам», равно как и по обыденным суждениям восприятия относительно людей, стоящих на полу, или о чашках весов, находящихся в равновесии.

Когда мы переходим к теориям высокого уровня, которые встречаются в точных науках, реакция людей будет отчасти иной. Простые люди не в состоянии «верифицировать» специальную теорию относительности. Действительно, в настоящее время обыкновенный человек даже не *учит* специальную теорию относительности, равно как и сравнительно элементарную математику, необходимую для ее понимания, хотя основы ее преподаются в ряде университетов в рамках начального курса физики. Простые люди полагаются на ученых в компетентной (и социально принятой) оценке подобного рода теории. Однако ученый, принимая во внимание нестабильность научных теорий, не отнесет даже такую успешную теорию, какой является специальная теория относительности, к числу теорий, «истинных» *tout court*². Однако мнение научного сообщества сво-

2. В общем и целом (фр.). — Прим. пер.

дится к тому, что специальная теория относительности «успешна» — в действительности, подобно квантовой электродинамике, она является беспрецедентно успешной, дающей «успешные предсказания» и «поддержанной широким набором экспериментов». И фактически на эти суждения полагаются другие члены общества. Различие между этим случаем и теми случаями институционализированных норм верификации, о которых речь шла выше, сводится, — кроме ни к чему не обязывающего прилагательного «истинный», — к особой роли экспертов и к институционализированному доверию, которое общество оказывает этим экспертам в случаях, когда речь идет о проверке теорий высокого уровня. Однако это — не более чем пример разделения интеллектуального труда (не говоря уже об отношениях интеллектуального авторитета) в обществе. Суждение, согласно которому специальная теория относительности и квантовая электродинамика являются «самыми успешными физическими теориями из числа тех, которыми мы обладаем» — это суждение, вынесенное авторитетными людьми, назначенными обществом, чей авторитетный статус закреплен в целом ряде практик и ритуалов и подобным образом институционализирован.

Недавно мне пришло в голову, что Витгенштейн вполне мог бы полагать, что вообще только *те* высказывания, которые могут быть верифицированы подобного рода «институционализированным» способом, могут быть истинными (или правильными, или корректными, или обоснованными). Тем самым я вовсе не хочу сказать, что когда-либо существовал философ, который придерживался той точки зрения, что *все* процедуры, которые в нашем обществе считаются «обоснованиями», и в самом деле являются таковыми. В общем и целом философы проводят различие между институтами, конституирующими наши понятия, и теми, которые имеют какой-то иной статус, хотя постоянно идут споры о том, как провести это различие. Я склонен предположить, что Витгенштейн думал о некотором подмножестве наших институционализированных норм верификации, определяющих, что в тех различных «языковых играх», в которые мы играем, говорить правильно, а что — нет, и полагал, что помимо этих норм не существует никаких объективных критериев правильности и неправильности. Хотя подобная интерпретация неплохо соответствует многому из того, о чем го-

ворит Витгенштейн — например, он подчеркивает необходимость «согласия в наших суждениях» для того, чтобы вообще иметь понятия, — я не уверен, что она является правильной. Ведь слишком неопределенно, кто такие эти «мы» в словах Витгенштейна о «наших суждениях»; и я не знаю, соответствуют ли его «формы жизни» упомянутым мною институционализированным нормам. Эта интерпретация, однако же, пришла мне в голову во время чтения его «Лекций и бесед». В этом произведении Витгенштейн отвергает как психоанализ, так и дарвиновскую теорию эволюции (хотя в отличие от позитивистов, он называл язык психоанализа *бессмысленным* и восхищался «одаренностью» Фрейда). Точка зрения Витгенштейна на психоанализ, который он называл «мифом», немного стоит, поскольку очень многие люди придерживались того мнения, — ошибочного, на мой взгляд, — что психоанализ так или иначе является бессмысленным. Однако его отрицание теории эволюции является довольно показательным³. Витгенштейн отталкивается от неблагоприятного для дарвиновской теории сравнения ее с физическими теориями: «Одна из наиболее важных особенностей

3. Касаясь теории эволюции, Витгенштейн говорил: «Люди строят свои убеждения на основаниях крайне ненадежных. Найдется ли хотя бы один, который скажет: «Я не знаю. Это интересная гипотеза, которое со временем может быть по-настоящему подтверждено?»» (*Wittgenstein L. Lectures on Aesthetics // L. W. Wittgenstein. Lectures and Conversations / Ed. Burrett C. Berkeley: University of California Press, 1967. P. 26*). Что означает быть «хорошо подтвержденной» в случае с теорией эволюции, Витгенштейн не объясняет, однако из соответствующего пассажа можно заключить, что он имеет в виду реальное *наблюдение* этого процесса («Видел ли кто-нибудь, что происходит этот процесс? Нет. Видел ли кто-нибудь его теперь? Нет. Данные о разведении животных — это всего лишь капля в море»).

Позицию Витгенштейна показательно сопоставить с позицией Моно:

«Теория эволюции путем естественного отбора в том виде, в каком она была сформулирована Дарвином, требовала открытия генетики Менделя, которое, конечно же, было сделано. Это пример, причем наиболее важный пример того, что подразумевается под содержанием теории, под содержанием идеи... Хорошая теория или хорошая идея всегда будет шире и богаче, чем ее создатель мог предполагать в свое время. Теорию следовало бы оценивать именно с точки зрения этого развития, когда все больше и больше данных попадает в ее оборот, даже если было невозможно предсказать, что так много войдет в ее состав» (*Monod J. On the Molecular Theory of Evolution // Problems of Scientific Revolutions: Progress and Obstacles to Progress in the Science / Ed. Harre R. Oxford, 1975*).

объяснения состоит в том, что оно должно работать, должно снабжать нас предсказаниями. Физика связана с инженерным делом. Мост не должен рушиться»⁴. И он говорит, что убеждения людей построены «на крайне шатких основаниях». «Если вы уверены, что дело должно обстоять определенного рода образом, то в конце концов вы совсем забудете о самой проблеме верификации», — заключает Витгенштейн.

Опять-таки, мне представляется, что великая дискуссия о понятии «аналитичности», которая ведется с 50-х годов, связана с желанием философов найти объективное, *неоспоримое* основание для своих аргументов. Понятие «аналитичности», то есть доктрина истины в силу одного только значения подверглась критике потому, что оно была *переиспользована* философами. Но почему философы были склонны заявлять, что такое множество выражений, не являвшихся ни в каком разумном смысле ни правилами языка, ни следствиями этих правил, было аналитическим или «концептуально необходимым»? Ответ, как я полагаю, состоит в следующем: философы думали, что идея, согласно которой существует определенное множество *правил языка*, на основе *которых* можно установить, что является рациональным, а что нет, обладала двумя преимуществами: (1) что «правила языка» представляют собой конститутивно институционализированные практики (или нормы, лежащие в основе подобного рода практик) и в качестве таковых обладают общественным статусом, который я описал; (2) в то же самое время утверждалось, что только философы (а не лингвисты) могли бы раскрыть подобного рода тайнства. Это была прекрасная идея, пока она была жива, однако ей суждено было погибнуть, и она погибла.

Любую концепцию, согласно которой существуют институционализированные нормы, определяющие, что является рационально приемлемым, а что — нет, я буду называть *критериальной* концепцией рациональности. Логические позитивисты, Витгенштейн (по крайней мере при той допустимо неопределенной интерпретации, что была предложена мной) и некоторые, хотя и не все философы «обыденного языка»⁵ из Оксфорда разделя-

4. Wittgenstein L. Lectures on Aesthetics // L. W. Wittgenstein. Lectures and Conversations / Ed. Burrett C. Berkeley: University of California Press, 1967. P. 25.

5. Можно было бы создать такую философию «обыденного языка», которая не была бы связана с общественной и «критериальной» верификацией фило-

ли критериальную концепцию рациональности, хотя они и расходились в других вопросах, таких, как, например, вопрос о том, считать ли не поддающиеся верификации высказывания «бессмысленными», или же вопрос о том, следует ли считать некоторые этические пропозиции «концептуально необходимыми».

Гамбит, на который я сослался в самом начале этой главы и который опровергает принцип верификации логических позитивистов, является *веским* как раз потому, что он опровергает любую попытку защищать критериальную концепцию рацио-

софских тезисов, если бы имелась возможность развить и поддержать концепцию, в рамках которой нормы, посредством которых осуществляется управление языковыми практиками, сами не поддавались бы открытию при помощи обычного эмпирического исследования. В книге «Должны ли мы понимать то, что мы говорим?» («Must We Mean What We Say») Стэнли Кэвел (Cavell) предпринял важный шаг в этом направлении, утверждая, что подобные нормы можно познать при помощи своего рода «знания о себе», которое он сравнил с озарением, достигаемым благодаря терапии, а также с трансцендентальным знанием, которое пытается найти феноменология. Хотя я и согласен с Кэвелом в том, что мое знание, будучи знанием человека, говорящего на родном для него языке, того, что определенные способы использования языка отклоняются от общепринятой нормы, а другие нет, не является «внешним» индуктивным знанием, — например, я могу знать, *без ссылки на имеющиеся факты*, что на использовании мной диалекте английского языка говорят о мышах как о «mice», а не как о «mouses», — я, тем не менее, склонен полагать, что этот факт привилегированного доступа говорящего не поддается распространению на *обобщения* по поводу правильности и неправильности словоупотребления. Если я говорю (а именно так и поступает Кэвел), что неотъемлемой частью правила, определяющего правильное использование утверждения формы «Х является произвольным», должно было бы являться то обстоятельство, что Х почему-то не внушает доверия, то я тем самым выдвигаю *теорию*, призванную объяснить мои интуитивные представления об отдельных случаях, а не просто излагаю эти интуитивные представления. Нечто подобное и в самом деле происходит в психотерапии, однако я не склонен наделять знание о себе какого угодно рода иммунитетом от критики со стороны других лиц, включая критику, которая в любом случае зависит от выдвижения *альтернативных* объяснений. Если же правомочность подобной критики признается, то деятельность, направленная на открытие подобного рода норм начинает походить на социальную науку или на историческое исследование, а в этих областях, как я утверждал, традиционные подходы, связанные с «научным методом», являются малопродуктивными. Об этом см. мою книгу: Putnam H. Meaning and Moral Sciences. L.; N. Y.: Routledge and Kegan Paul, 1978.

В любом случае, каким бы ни был статус подобных норм, определяющих использование языка, я не вижу никаких оснований считать, что они определяют объем таких выражений, как «рационально приемлемый», «обоснованный», «в достаточной степени подтвержденный» и т. п.

нальности, то есть потому, что он опровергает тезис, согласно которому все, что является рационально верифицируемым, должно быть критериально верифицируемым.

Суть моей позиции заключается в следующем: хотя философы, о которых я упоминал, часто рассуждали так, как будто их доказательства имели ту же степень неопровержимости, которой располагают математическое доказательство и физический эксперимент; хотя логические позитивисты называли свою деятельность *логикой* науки; хотя витгенштейнианцы проявили необычайное высокомерие по отношению ко всем тем философам, которые были не в состоянии «понять», что любая философская деятельность, имевшая место до появления витгенштейнианства и продолжающая иметь место после его появления, лишена всякого смысла; хотя философы обыденного языка ссылались на аргументы друг друга и на аргументы философов, не являвшихся приверженцами философии обыденного языка, как на «грубые ошибки» (как будто философские ошибки подобны ошибкам, совершаемым при решении арифметических задач), тем не менее ни одна из этих философских позиций не может быть верифицирована тем убедительным и культурно общепризнанным способом, который я описал. Короче говоря, если верно, что только те высказывания, которые могут быть критериально верифицированы, могут быть рационально приняты, само высказывание, содержащее эту мысль, не может быть верифицировано критериально и, следовательно, не может быть рационально приемлемым. Если вообще существует такая вещь, как рациональность, и мы, участвуя в деятельности, связанной с *речью* и *аргументацией*, связываем себя с верой в *некоторое* понятие рациональности, то оказывается самоопровергающейся всякая защита позиции, согласно которой *рациональность тождественна тому, что институционализированные нормы культуры определяют в качестве ее частных случаев, или же что она находит в этих случаях свое воплощение*. Ведь опираясь исключительно на эти нормы, нельзя будет удостовериться в том, что подобного рода точка зрения является правильной или хотя бы вероятно правильной.

Я вовсе не считаю, что рациональная аргументация или рациональное обоснование невозможны в философии. Скорее я пришел к необходимости признать то, что, по-видимому, оче-

видно если не для философов, то для простых людей, а именно то, что мы не можем обращаться к *общественным* нормам для того, чтобы решить, что рационально доказано и обосновано в философии, а что — нет. Заявление, которое все еще часто приходится слышать, что философия — это «концептуальный анализ» и что *понятия сами* определяют, какие философские аргументы правильны, представляет собой, коль скоро оно объединяется с учением, согласно которому понятия представляют собой нормы или правила, лежащие в основе *общественных* языковых практик, скрытую форму утверждения, согласно которому всякое рациональное обоснование в философии — это критериальное обоснование, а философские истины в той же степени *поддаются общественно значимому доказательству*, что и научные истины. Такая точка зрения кажется мне попросту неразумной в свете всей истории предмета, не исключая и новейшую историю.

То, что имеет силу в отношении философской аргументации, в равной степени имеет силу и в отношении аргументов, связанных с религией и со светской идеологией. Спор между интеллектуалом-либералом и интеллектуалом-марксистом будет в конечном счете носить тот же самый характер, что и философский спор, даже если в рамках этого идеологического спора будет использовано больше эмпирических фактов. У всех у нас имеются религиозные, политические или философские взгляды и убеждения, и все мы защищаем их и критикуем доводы других. В самом деле, даже в «науке», если не ограничивать ее рамками точных наук, — то есть в истории, социологии, клинической психологии, — мы используем доводы точно такого же типа. Верно, что логические позитивисты расширили данное ими описание «научного метода» для того, чтобы включить в него эти области знания, однако при таком расширении невозможно *ясно* показать, исключает ли этот метод вообще что-либо (смотри главу 8).

Мне напомнят, что согласно точке зрения позитивистов, принцип верификации является «познавательльно бессмысленным»⁶. Они говорили, что это *заявка* и как таковая она не мо-

6. Суть наиболее слабого аргумента, выдвинутого в защиту принципа верификации, истолкованному как заявка, состояла в том, что этот принцип «экс-

жет быть ни истинной, ни ложной. Однако они приводили доводы в защиту своей заявки, и их аргументы были неудачными, — и они были обречены быть таковыми. Так что суть моей критики сохраняет свою силу.

В общем, логические и позитивисты и Витгенштейн (и возможно, так же и поздний Куайн) создали философские концепции, которые не оставляют места для рациональной деятельности в философии. Именно поэтому их воззрения являются самоопровергающимися. Поэтому точно так же тот небольшой гамбит, который я обсуждал в начале этой главы, представляет собой действительно важный аргумент этого рода, который философы называют «трансцендентальным аргументом». Суть его заключается в том, что рассуждение по вопросу о природе рациональности (занятие для философов *par excellence*) представляет собой деятельность, предполагающую понятие рационального обосно-

плицировал» «доаналитическое» понятие осмысленности. (Обсуждение этого тезиса см. в моей статье «Как нельзя говорить о значении»: Putnam H. How Not to Talk about Meaning // Putnam H. Mind, Language and Reality. Philosophical Papers. Vol. 2. Cambridge: Cambridge University Press, 1975). Рейхенбах защищал принцип верификации (Reichenbach H. Experience and Prediction. Chicago, 1938) в форме, сохраняющей все различия в значении, релевантные поведению. На явным образом напрашивающееся возражение, что неэмпирическая вера в божественное (Рейхенбах использовал пример почитания кошек в Древнем Египте) может изменить поведение, Рейхенбах отвечал предложением перефразировать высказывание «кошки — божественные животные» как «кошки вызывают чувство благоговения у почитателей кошек». Само собою понятно, что принятие подобного рода замены не оставило бы неизменным поведение почитателей кошек!

Наиболее интересная идея была предложена Карнапом. Согласно Карнапу, все рациональные реконструкции являются заявками. Единственное, с чем связаны проблемы фактического плана, — это логические и эмпирические последствия принятия той или иной реконструкции. (Карнап сравнивал «выбор» рациональной реконструкции с выбором двигателя для самолета). Он пришел к выводу, что в философии следовало бы проявлять терпимость к различным рациональным реконструкциям. Тем не менее этот принцип терпимости, как его называет Карнап, предполагает принцип верификации. Поскольку доктрина, согласно которой ни одна рациональная реконструкция не является однозначно правильной и не соотносится с тем способом, которым вещи «в действительности существуют», доктрина, согласно которой все «внешние вопросы» не имеют познавательного смысла, — это по сути дела и есть принцип верификации. Применение же принципа терпимости к принципу верификации привело бы к логической ошибке, которую обычно называют «кругом в определении».

вания более широкое, нежели то, что было предложено позитивистами, и конечно же, более широкое, чем институционализируемая критериальная рациональность.

АНАРХИЗМ ЯВЛЯЕТСЯ САМООПРОВЕРГАЮЩИМСЯ

Позвольте мне теперь обсудить другую, весьма отличную от описанной выше тенденцию в философии. Огромное число читателей было очаровано, а большинство философов науки — напугано книгой Томаса Куна «Структура научных революций»⁷ благодаря тому акценту, который был в ней сделан на *иррациональных*, как казалось, факторах, обуславливающих принятие научной теории, а так же благодаря использованию в ней таких терминов, как «конверсия» и «переключение гештальта». В самом деле, Кун сделал ряд важных замечаний о научных теориях и о том, как следовало бы рассматривать научную деятельность. Я уже выражал свою уверенность в том, что такие понятия, как «парадигма», «нормальная наука», «научная революция» имеют важное значение, однако в данном случае я хочу сосредоточиться на тех пунктах книги Куна, которым я *не* симпатизирую, на том, что я ранее описал как «крайний релятивизм Куна».

Более проницательных читателей увлекло такое прочтение книги Куна, согласно которому, по его словам, в науке нет места *рациональному* обоснованию, а происходят *лишь* переключения гештальта и конверсии. Кун отверг эту интерпретацию «Структуры научных революций», и ввел понятие «непарадигматической рациональности», которое тесно связано с понятием, названным мною «некритериальной рациональностью»; быть может, эти два понятия даже являются тождественными.

Тенденция, которую большинство читателей обнаружило, как они полагали, в «Структуре научных революций» Куна, со всей определенностью проявилась в работе Пола Фейерабенда «Против метода». Подобно Куну, Фейерабенд особое внимание обращается на те способы, посредством которых различные культу-

7. Kuhn T. The Structure of Scientific Revolutions. Chicago, 1962. [Имеется русский перевод: Кун Т. Структура научных революций / Пер. с англ. И. З. Налецова. М.: Прогресс. 1975. — Прим. пер.].

ры и исторические эпохи производят различные парадигмы рациональности. Он считает, что *наши* понятия научной рациональности в значительной степени определяются факторами, которые мы назвали бы иррациональными. По сути дела он утверждает, хотя и не говорит этого буквально, что современная научно-технологическая концепция рациональности в соответствии со своими собственными стандартами является мошенничеством. (Мне кажется, что подобного рода мотив я обнаружил у Мишеля Фуко). Причем Фейерабенд идет дальше Куна или Фуко, допуская, что даже хваленое инструментальное превосходство науки может быть не чем иным, как мистификацией. Фейерабенд утверждает, что знахари, лечащие при помощи молитв, могут больше преуспеть в облегчении нашей боли, нежели профессиональные врачи.

Однако я намереваюсь подвергнуть обсуждению вовсе не эти ужасающе радикальные заявления, хотя они и позволяют Фейерабенду именовать свою позицию «анархизмом». Я намереваюсь обсудить заявление, сделанное Куном в «Структуре научных революций» и последующих статьях, а Фейерабендом — в книге «Против метода» и статьях, носящих технический характер. Это тезис о *несоизмеримости*. Я хочу показать, что этот тезис, равно как и тезис логического позитивизма о значении и верификации, является самоопровергающимся. Короче говоря, я намереваюсь заявить, что обе эти две наиболее влиятельные философии науки XX века, философии, которые, собственно говоря, и интересовали ученых и нефилософов как таковых, и, по всей видимости, единственные философии, о которых образованный читатель вообще когда-либо слышал, являются самоопровергающимися. Конечно же, будучи философом науки, я испытываю некоторое беспокойство по поводу подобного рода ситуации. Мы вскоре столкнемся с вопросом, что со всем этим делать.

Тезис о несоизмеримости — это тезис о том, что термины, используемые в другой культуре, например, термин «температура» в том виде, в каком он использовался ученым XVII века, не совпадают по своему смысловому значению или же по своей референции с теми терминами и выражениями, которые находятся в *нашем* распоряжении. Как подчеркивает Кун, ученые, работающие в различных парадигмах, живут в «различных ми-

рах». Термин «электрон» в том виде, в каком он употреблялся в 1900 году, относился к объекту одного «мира»; в своем же нынешнем употреблении он относится к совершенно иному «миру». При этом предполагается, что тезис о несоизмеримости касается как языка наблюдений, так и «теоретического языка». Согласно Фейерабенду, обыденный язык — *это* просто ложная теория.

Здесь сразу же напрашивается возражение: если бы тезис о несоизмеримости и в самом деле был истинным, то мы вообще были бы не в состоянии переводить с других языков; даже наш собственный язык в том виде, в каком он существовал в прошлом, не поддавался бы переводу. А если мы совсем не в состоянии интерпретировать шумы, издаваемые организмами, то у нас нет никаких оснований для того, чтобы считать их *мыслящими, говорящими существами* или даже *личностями*. Короче говоря, если бы Фейерабенд (и Кун в его подчеркивании несоизмеримости) были бы правы, то представители других культур, включая ученых XVII века, могли бы быть поняты нами только как животные, издающие — в виде реакции на стимулы — шумы, отчасти весьма странным образом походящие на английский и итальянский языки. Было бы полным противоречием сообщать нам, что Галилей имел в своем распоряжении понятия, «несоизмеримые» с нашими, *а затем продолжать подробно их описывать*.

Эта проблема была поставлена Смартom в статье, посвященной взглядам Фейерабенда и написанной в весьма благожелательном для последнего духе:

«К числу нейтральных фактов мы можем, конечно, отнести тот факт, что для того, чтобы увидеть Меркурий, мы должны направить телескоп, скажем, поверх данного дерева, а не поверх данной дымовой трубы, как это предсказывает ньютоновская теория. Несомненно, можно говорить о деревьях, дымовых трубах и телескопах таким способом, который не зависит от выбора между теориями Ньютона и Эйнштейна. Однако на это Фейерабенд мог бы заметить, что для построения нашего телескопа мы используем евклидову геометрию и нерелятивистскую оптику. И хотя они не дают полной истины относительно телескопа, упомянутых деревьев и дымовой трубы, а также планеты Меркурий, тем не менее для того, чтобы обсудить эмпирическую проверку общей теории относительности, разумно описы-

вать явления именно таким способом, потому что у нас есть теоретические основания считать, что наши предсказания останутся без изменения в тех случаях, когда мы пользуемся этим средством вычисления»⁸.

Однако со спасительным ходом Смарта связана определенная проблема, суть которой заключается в том, что я должен *хотя бы отчасти* понимать евклидов нерелятивистский язык для того, чтобы утверждать, что «предсказания» — те же самые. Если *каждое слово имеет иное значение*, то в каком смысле можно утверждать, что предсказания остаются «без изменений»? Каким образом я вообще могу переводить логические частицы (слова «если..., то...», «не» и т. д.) на итальянский язык XVII века, если я не могу найти руководство для перевода, связывающее итальянский язык XVII века и современный английский язык и делающее систематически осмысленным текст XVII века, причем как сам по себе, так и по отношению к тому, что он описывает? Даже если я рассуждаю, используя обе теории (как это и предусматривает Сمارт), то откуда я могу почерпнуть уверенность в том, что какое-либо слово из ньютоновской теории эквивалентно какому-либо слову из теории относительности?

Проблема, с которой мне приходится иметь дело, приобретает еще более четкие очертания, если мы сопоставим с ней некоторые наблюдения Куайна и Дэвидсона, касающиеся языкового значения и практики перевода. Если предположить, что мы можем найти схему перевода, которая позволяет иметь дело с текстом XVII века, по крайней мере в контексте, определяемом нашими интересами, и с учетом той цели, для которой нам нужен перевод, то какой смысл имеет *в данном контексте* заявление о том, что перевод «в действительности» не ухватывает смысл или референцию оригинала? В конце концов, у нас нет и, судя по всему, не может быть критериев тождественности смысла и референции вне наших схем перевода и явно либо неявно сформулированных требований к эмпирической адекватности перевода. Утверждение, согласно которому в переводе не

8. Smart J. J. Conflicting Views about Explanation // Boston Studies in the Philosophy of Science / Ed. Cohen R., Wartofsky M. Vol. 2. N. Y.: Humanities Press, 1965. [Ср. русский перевод: Смарт Дж. Конфликтующие точки зрения по проблеме объяснения / Пер. с англ. А. Л. Никифорова // Структура и развитие науки / Под ред. Грязнова Б. С., Садовского В. Н. М.: Прогресс, 1978. С. 346].

схвачен точный смысл или значение оригинала, может быть понято как допущение, что возможна более совершенная схема перевода. Однако утверждение, что любая возможная схема перевода не в состоянии схватить «подлинный» смысл или референцию оригинала, выглядит осмысленным только на первый взгляд. Синонимия существует только в качестве отношения, скорее даже в качестве семейства отношений, используемых нами для того, чтобы сравнивать различные выражения в целях интерпретации. Идея о том, что существует что-то подобное «реальной» синонимии, независимой от всех рабочих практик взаимной интерпретации, была развенчана как миф⁹.

Предположим, что кто-то говорит нам, что немецкое слово «Rad» можно перевести как «колесо». Если он дополнительно сообщает нам, что его перевод несовершенен, то мы, естественно, ожидаем от него, что он нам укажет, как можно было усовершенствовать его перевод, снабдить каким-нибудь истолкованием и т. д. Однако если этот человек продолжает настаивать на том, что слово «Rad» можно перевести как «колесо», но в действительности оно не относится к колесам, или даже к любому объекту, упомянутому в нашей концептуальной схеме, то что мы почерпнем из всего этого? Сказать, что слово А можно перевести как «колесо», или еще как-то, означает утверждать, что в той степени, в какой на перевод можно положиться, А *указывает* на колеса.

Вероятно, причина того, что тезис о несоизмеримости столь сильно увлекает людей, заключается, помимо той привлекательности, которая, по-видимому, свойственна всем противоречивым идеям, в тенденции смешивать понятие (*concept*) и концепцию (*conception*). В той степени, в какой смутным является различие между аналитическими и синтетическими суждениями, смутно и различие между понятием и концепцией; тем не ме-

9. Судя по всему, Патнэм имеет в виду статьи Нельсона Гудмена, посвященные проблеме синонимии, в которых было подвергнуто критике неконтекстуальное, асбютистское понимание этого семантического отношения и предложена оригинальная, «ослабленная» версия интерпретации данного отношения. Об этом см.: Гудмен Н. О сходстве значения / Пер. с англ. Т. А. Дмитриева // Гудмен Н. Способы создания миров. М.: Идея-Пресс; Праксис, 2001. С. 339–350; Гудмен Н. О некоторых различиях, касающихся значения / Пер. с англ. Т. А. Дмитриева // Там же. С. 351–360. — Прим. пер.

нее все интерпретации подразумевают это различие, даже если оно касается только самой интерпретации. Когда мы переводим некоторое слово, например, «температура», мы приравниваем референцию и в той степени, в какой мы держимся данного перевода, смысл переводимого термина «температура» к смыслу нашего слова «температура», по крайней мере к тому смыслу, в котором оно используется в данном контексте. (Разумеется, существуют различные средства, такие, как специальные интерпретации, которые мы можем использовать для того, чтобы очертить пределы или размах использования нами слова «температура» или какого-либо иного возможного слова в этом контексте). В этом случае мы приравниваем то «понятие», о котором идет речь, к нашему собственному «понятию» температуры. Однако подобного рода процедура вполне может быть согласована с тем фактом, что ученые XVII века могли обладать представлением о температуре, отличным от того, которое имеется в нашем распоряжении, то есть иным рядом убеждений (*beliefs*) о ней и ее природе, иными «образами знания», а также иными предельными убеждениями по поводу множества других вещей. Из того, что концепции различаются, вовсе не следует, как это принято часто предполагать, что «на самом деле правильный» перевод вообще невозможен. Напротив, мы не могли бы утверждать, что концепции различаются, и определить, чем же, собственно говоря, они отличаются, если бы мы не могли переводить.

Можно, однако же, поинтересоваться, откуда мы знаем, что схема перевода «работает», если концепции всегда оказываются различными? Ответ на этот вопрос, который дают многие мыслители от Вико до современных философов, заключается в том, что успех интерпретации не требует того, чтобы переводимые убеждения были *тождественными* нашим собственным; в действительности он требует только того, чтобы они были *понятными* для нас. Это — основа всех разнообразных максим благожелательного доверия к интерпретации или «пользы сомнения», таких, как «истолковывайте их слова таким образом, чтобы они предстали верующими в истину и почитающими благо» или «истолковывайте их так, чтобы их убеждения оказались разумными с точки зрения того, что они учили и с какого рода опытом им приходилось иметь дело»; сюда же относится выс-

казывавшееся самим Вико пожелание приписывать тем людям, чьи рассуждения и поступки мы интерпретируем, максимальную степень *человечности*. Тот факт, что мы, с точки зрения универсального человеческого опыта, способны *достичь* подобного понимания, относится к числу определяющих фактов человеческого опыта, развертывающегося в мире различных культур, взаимодействующих друг с другом в ходе исторического процесса и подвергающихся по отдельности более или менее быстрым изменениям. Это означает, что мы способны *осмысленно* истолковывать убеждения, желания и изречения друг друга.

Не удивительно, что Кун и Фейерабенд отвергают любую идею конвергенции в научном знании. Поскольку мы говорим не о тех же самых предметах, о которых говорили ученые прошлого, то мы и не приобретаем все больше знания относительно тех же самых макроскопических и микроскопических объектов. Кун утверждает, что научный «прогресс» носит исключительно инструментальный характер: например, мы получаем возможность быстрее перемещать людей из одного места в другое и т. д. Однако эта позиция чересчур непоследовательна. Если такой оборот речи, как «перевозка людей с одного места на другое» не сохранит определенную степень устойчивой референции, то как мы сможем сколько-нибудь определенно постичь понятие инструментального успеха?

Аргумент, который я только что использовал, самым тесным и непосредственным образом связан со знаменитым аргументом Канта относительно предварительных условий эмпирического знания. Отвергая довод о том, что будущее может совершенно выйти за пределы сферы действия законов, может опровергнуть любое индуктивное заключение, сделанное нами, Кант указывал, что если вообще существует какое-либо будущее, — по крайней мере какое-либо будущее *для нас*, по крайней мере какое-либо будущее, которое мы, как мыслящие существа, можем постичь и поместить в нашу понятийную систему координат для того, чтобы узнать, являются ли наши предсказания истинными или ложными, то в этом случае многие закономерности *не* должны нарушаться. В противном случае, почему мы называем это *будущим*? Например, когда мы представляем шары, появляющиеся из урны в некоторой «нерегулярной» последовательности, мы забываем о том, что мы просто-напросто *были бы не в состоянии*

утверждать, что перед нами находятся шары или что они появляются в беспорядке, если бы мы не находились в зависимости от многих закономерностей. Сравнение предполагает, что существует какая-то степень соизмеримости.

В ответ на эту критику Кун и Фейерабенд могли бы возразить, хотя и не совсем охотно, что существует определенного рода дихотомия между наблюдаемым и теоретическим. Они могли бы предположить, что соизмеримость, переводимость и даже конвергенция относятся к языку наблюдения, и ограничить область действия тезиса о несоизмеримости теоретическим словарем. Однако и в этом случае возникли бы проблемы (почему бы нам в духе Рамсея не описать значения теоретических терминов через их отношение к языку наблюдения?). Кун и Фейерабенд, однако же, с полным правом отвергают эту альтернативу, поскольку в действительности принципы доброжелательности необходимы как при интерпретации «языка наблюдения», так и при интерпретации «языка теории». Возьмем, для примера, общеизвестное слово «трава». Различные говорящие, в зависимости от места и времени своего проживания располагают различными образцами чувственного восприятия травы (травы имеют различные цвета и различный размер в разных местах) и различными представлениями о траве. *Даже если все говорящие должны знать, – при том, что иное представление их не устраивало бы, – что трава – это растение, представления о растении в настоящее время включают в себя идею фотосинтеза, тогда как двести лет назад представление о растении не включало в себя подобного рода идеи.* Без интерпретативной доброжелательности, которая позволяет нам, – по крайней мере в обычных контекстах, – приравнять слово «растение», употребленное двести лет назад, и слово «растение», произнесенное сегодня, а также слово «трава», употребленное двести лет назад, и сегодняшнее слово «трава», было бы невозможно высказать что-либо о той референции слова «трава», которым оно обладало двести лет назад. Причем от принципа доброжелательности зависит не только интерпретация слов, обозначающих естественные виды; слово «хлеб» в действительности породило бы те же самые трудности. В самом деле, без благожелательности, проявляемой в процессе интерпретации, мы вряд ли могли бы приравнивать референции такого используемого для обозначения цвета термина, как «крас-

ный», в ситуации его произнесения многими людьми. Мы всегда интерпретируем дискурс как целое; поэтому интерпретация терминов наблюдения зависит от интерпретации теоретических терминов, тогда как интерпретация теоретических терминов зависит от интерпретации терминов наблюдения.

То, что я получил, — это опять-таки трансцендентальный аргумент. В силу наших основополагающих представлений мы вынуждены считать *личностями* не только себя самих в том виде, в каком мы существуем в настоящее время; мы вынуждены относить к числу личностей также прошлые состояния нашего Я, своих предков и представителей других культур в прошлом и настоящем; *а это, как я утверждал ранее, означает наделение их общей референцией и общими понятиями, как бы ни отличались представления, которые мы им тоже приписываем. Притом мы не только имеем дело с общими объектами и понятиями, но в той степени, в какой наши усилия, направленные на интерпретацию, оказываются успешными, мы обладаем также общими представлениями о том, что разумно, естественно и т. д.* Ибо вспомним, что исчерпывающее обоснование схемы интерпретации это такое обоснование, которое делает поведение других людей по крайней мере минимально разумным с *нашей* точки зрения. Как бы ни различались наши образы знания и наши концепции рациональности, мы обладаем громадным запасом допущений и убеждений по поводу того, что является разумным, причем разделяем их даже с наиболее причудливой культурой, которую сумели успешно истолковать.

ПОЧЕМУ РЕЛЯТИВИЗМ ЯВЛЯЕТСЯ ВНУТРЕННЕ ПРОТИВОРЕЧИВЫМ?

Та истина, что (последовательный) релятивизм является внутренне противоречивым, для философов давно уже банальный трюизм. Помимо всего прочего, разве не *само собою разумеется*, что противоречиво *придерживаться* какой-то точки зрения, но в то же самое время считать, что *ни одна* точка зрения не является более оправданной или правильной по сравнению с другой? Алан Гарфинкель весьма остроумно сформулировал эту точку зрения. Разговаривая со своими калифорнийскими студентами, он, подражая их манере речи, однажды сказал: «У нас с вами

могут быть разные позиции, но я знаю, что с *моей точки зрения* релятивизм является *ошибкой*»... Если любая точка зрения хороша, то почему тогда *точке зрения, согласно которой релятивизм является ложным*, не быть столь же правильной, сколь и любой другой?

Изобилие распространяемых сегодня релятивистских учений (причем распространяемых весьма толковыми мыслителями) указывает на то, что такого простого опровержения недостаточно. Почему разумный релятивист должен был бы заключить, что всякое мнение (*для него*) столь же истинно, как и любое другое? Он не может воспрепятствовать вам (или Алану Гарфинкелю) заявить, что его, — релятивиста, — точка зрения *для вас не является истинной* или оправданной; но если релятивист будет начеку, то он может остроумно ответить, что истина для вас *для него* значит намного меньше, чем истина для него самого. В конце концов, разве ваши *собственные* понятия не являются для вас самыми важными? Будет ли в таком случае противоречием считать «истинное», «обоснованное» и тому подобные понятия *понятиями относительными*?

Ответ заключается в том, что подобная рода трактовка все-таки является противоречивой, но в то же самое время она *действительно* требует более тщательно разработанного доказательства, нежели (весьма симпатичная) шутка, придуманная Гарфинкелем. Важно отметить, что если все является относительным, то и относительное тоже относительно. Однако это замечание заслуживает объяснения!

По всей видимости, Платон был первым философом, использовавшим, — против Протагора, — тот аргумент, который я имею в виду. Протагор, который, судя по всему, был релятивистом до мозга костей, утверждал, что, когда я говорю X, мне в действительности следовало бы говорить: «Я *думаю*, что X». Например, по поводу моего утверждения «снег бел» Протагор сказал бы, что я в действительности имею в виду то, что Хилари Патнэм думает, что снег бел, а когда соответствующее утверждение делает Роберт Нозик, то это означает, что он думает, что снег бел. В более утонченной форме эту идею можно было бы выразить следующим образом: когда я говорю: «Снег бел», я использую это утверждение для того, чтобы заявить, что *снег бел* истинно для меня, тогда как Роберт Нозик обычно использует те же са-

мые слова для того, чтобы заявить, что *снег бел* истинно для него (или по крайней мере он бы считал, что его высказывание является истинным только в том случае, когда оно оказывается истинным для *него самого*). Отсюда, согласно точке зрения Протагора, следует, что ни одно утверждение не имеет одинакового значения (*meaning*) для меня и для кого либо еще; как мы уже имели возможность убедиться прежде, существует тесная связь между релятивизмом и тезисом о несоизмеримости. Контраргумент, выдвинутый Платоном, состоял в следующем: если высказывание *X* означает «я думаю, что *X*», то, согласно точке зрения Протагора, мне в действительности следовало бы выражаться так:

(1) Я думаю, что я думаю, что снег бел.

Но процесс добавления выражения «я мыслю» может носить бесконечный характер! Согласно точке зрения Протагора, в итоге значение выражения «снег бел» приобрело бы (в отличии от исходной формы (1)), следующий вид:

(2) Я думаю, что я думаю, что я думаю, что я ... (и так до бесконечности), что снег бел.

Такого рода исход Платон расценивает как *reductio ad absurdum*¹⁰. Однако аргумент Платона не столь хорош, как кажется на первый взгляд. Отчего бы Протагору не согласиться, что его анализ применим к нему самому? Отсюда вовсе не следует, что он *должен* быть применим к нему самому бесконечное число раз, но только то, что он может быть применен к нему *какое-то конечное число раз*. Однако Платон подметил одно очень важное обстоятельство.

Когда мы впервые сталкиваемся с релятивизмом, то идея *кажется* очень простой. Идея, если отталкиваться от ее первой естественной формулировки, состоит в том, что каждая личность (или, если следовать современной «социологической» формуле, всякая культура или всякое «рассуждение») обладает своими собственными точками зрения, стандартами, предпосылками, а истина (а также обоснование) от *них* зависят. При этом считается чем-то само собою разумеющимся, что вопрос о

10. Сведение к абсурду (лат.). — Прим. пер.

том, является ли X истинным (или обоснованным) относительно этих взглядов, стандартов и предпосылок, сам по себе носит «абсолютный» характер.

Современные структуралисты вроде Фуко высказываются в том духе, как если бы *соотнесенное с рассуждением* обоснование само было чем-то абсолютным, то есть вовсе не относительным. Однако если высказывания формы « X является истинным (обоснованным) относительно личности P » сами являются истинными или ложными *абсолютно*, то всё же *существует* абсолютное понятие истины (или обоснования), а не только истины-для-меня, истины-для-Нозика, истины-для-вас и т. д. *Последовательный* релятивист должен был бы утверждать, что вопрос о том, является ли X истинным или ложным *относительно P* , сам тоже является относительным. Как заметил Платон, в этом пункте наше понимание того, что собой представляет точка зрения релятивизма, начинает колебаться.

Платоновская форма критики релятивизма в наши дни не имеет широкого использования. Однако она нашла себе прекрасное продолжение в аргументе Витгенштейна относительно приватного языка, о котором шла речь в третьей главе.

Большинство комментаторов считает, что аргумент относительно приватного языка направлен против теории истины как «копии». Превосходное доказательство Витгенштейном того, что теория референции как подобия не годится даже в том случае, когда речь идет об указании (*reference*) на ощущения, конечно же, является частью имеющей долгую историю критики метафизического реализма. Однако я склонен рассматривать этот аргумент в качестве двух совершенно традиционных аргументов (по крайней мере Кант одобрил бы оба!), направленных как против реалистической, так и против релятивистской позиций. Попытка же истолковать весь аргумент в целом как антиреалистический может выглядеть весьма натянутой.

Та форма релятивизма, которую намеревался подвергнуть критике Витгенштейн, известна как «методологический солипсизм». «Методологический солипсист» — это не-реалист или «верификационист», который согласен с тем, что истину следует считать в известной степени зависящей от рациональной приемлемости, но который в то же самое время считает, что всякое обоснование в конечном счете осуществляется в терминах,

выражающих опыт, знание о котором, в свою очередь, для каждого из нас является *индивидуальным*. Таким образом, я обладаю *своим* знанием о том, какого рода данные опыта, имеющиеся в *моем* распоряжении, верифицировали бы высказывание, что снег бел, а Боб Нозик обладает *своим* знанием о том, какого рода данные опыта, имеющиеся в *его* распоряжении, верифицировали бы высказывание, что снег бел. Это означает, что любое высказывание обладает различным смыслом для разных мыслителей.

Аргумент Витгенштейна представляется мне превосходным аргументом, направленным против релятивизма в целом. Суть аргумента состоит в том, что релятивист в конечном счете не в состоянии провести осмысленное различие между тем, что является *правильным* (*being right*) и тем, что *кто-то считает правильным* (*thinking he is right*); а это в конечном счете означает, что нет никакого различия между *актом утверждения* или *мышления*, с одной стороны и *произнесением звуков* (или *продуцированием ментальных образов*), с другой. Но это, в свою очередь, в соответствии с логикой данной концепции, означает, что я по сути дела являюсь не *мыслителем*, а *просто животным*. Разделяя подобное мнение, релятивист совершает своего рода ментальное самоубийство.

Для того, чтобы понять правоту Витгенштейна, давайте обратимся к рассмотрению проблемы, о которой у Витгенштейна речи не шло, а именно проблемы того, как релятивист мог бы *попытаться* провести то различие, на которое он, по мнению Витгенштейна, был не способен, — различие между тем, что является *правильным*, и тем, что *некто считает правильным*.

Релятивист мог бы воспользоваться идеей, согласно которой истина представляет собой *идеализацию* рациональной приемлемости. Он мог бы предположить, что *X истинно-для-меня*, если «*X обосновано для меня*» было бы истинно, при условии, что я достаточно внимательно наблюдал, достаточно долго размышлял и т. д. Однако условные высказывания в *сослагательном наклонении* формы «*Если бы я был должен..., то я бы думал так*», подобно всем прочим высказываниям, разными философами интерпретируются по-разному.

Метафизический реалист может интерпретировать высказывания о том, что «*имело бы место, если*» как истинные или лож-

ные в абсолютном смысле этих слов, вне зависимости от того, будем ли мы правы, принимая или отрицая эти высказывания. Если релятивист истолковывает высказывания о том, во что бы он *верил* при наличии определенных условий в подобного рода реалистическом духе, то тем самым он уже признает *один* класс абсолютных истин, и потому перестает быть релятивистом.

Не-реалист или «внутренний» реалист истолковывает условные высказывания как высказывания, которые мы, — подобно всем другим высказываниям, — понимаем по большей части благодаря постижению условий их *обоснования*. Это вовсе не означает, что «внутренний» реалист *отбрасывает* различие между истиной и обоснованием; это означает всего-навсего то, что истина (*идеализированная* обоснованность) постигается нами тем же способом, что и прочие понятия, то есть при помощи по большей части неявно выраженного понимания факторов, которые делают утверждение о том, что нечто является истинным, рационально приемлемым. Может ли *релятивист* истолковывать высказывания о том, во что он *был бы* склонен верить при идеальных условиях в этом не-реалистическом или «внутренне» реалистическом духе?

Позвольте мне напомнить, что не-реалистическая позиция, изложенная мною в третьей главе, предполагает *объективное* понятие рациональной приемлемости. Антиреалист отвергает представление, согласно которому истина представляет собой соответствие с «уже готовым миром». Именно благодаря этой особенности своей позиции антиреалист и является *анти-метафизическим-реалистом*. Однако отказываться от метафизической теории истины как соответствия — это одно, а считать истину или рациональную приемлемость чем-то *субъективным* — это совсем другое. Нельсон Гудмен, который считает истину и рациональную приемлемость разновидностями более общего предиката «правильность», применимого как к произведениям искусства, так и к высказываниям, кратко выразил эту мысль следующим образом:

«Итак, истинность высказываний и правильность описаний, представлений, экзemplификаций, выражений — композиции, рисунка, дикции, ритма — это прежде всего вопрос соответствия (*fit*) тому, к чему осуществляется та или иная референция, или другим представлениям, или способам и методу организации.

Различия между соответствием версии миру, мира — версии и версий — друг другу исчезают, когда признана роль версий в создании миров, которым они соответствуют. Знание или понимание рассматриваются как переход от приобретения истинных убеждений к открытию и изобретению соответствия всех видов»¹¹.

В целом, однако же, *цель* релятивизма, его определяющая характеристика сводится к отрицанию сколько-нибудь вразумительного понятия *объективного* «соответствия». Поэтому релятивист не в состоянии уразуметь истолкования истины в терминах *объективных* условий обоснования.

Попытка использовать *условные высказывания* для того, чтобы выразить различие между тем, что *является правильным*, и тем, что *кто-то считает правильным*, терпит неудачу, поскольку релятивист не располагает *объективным* понятием правильности, которое он мог бы применить для оценки этих высказываний; та же самая ситуация имеет место и в отношении других видов высказываний.

Наконец, если современные релятивисты, подобно Протагору в древности, решают, стиснув зубы, настаивать на том, что не существует различия между «быть правым» и «считать себя правым», то есть что различие между тем, что является обоснованным, и тем, что человек считает обоснованным, неприменимо к *нему самому*, то о чем же, позвольте нас спросить, идет речь в этом случае, как не о простом издании звуков в надежде на то, что у говорящего возникнет *ощущение* правоты? Что же в этом случае будет представлять собой *мышление*, если не продуцирование образов и аналогов предложений в уме с надеждой на то, что у мыслящего появится субъективное ощущение правоты? В конечном счете релятивист должен прийти к отрицанию того, что мысли вообще *на* что-то указывают и *к* чему-то отсылают, причем неважно, будут ли эти процессы интерпретироваться в реалистическом или не-реалистическом духе; ведь он не в состоянии провести различие между мыслью о том, что чья-то мысль *на* что-то указывает, и самим актом мысли, направленным на

11. Goodman N. The Ways of Worldmaking. Hackett Publishing Company, 1978. P. 138. [Ср.: Гудмен Н. Способы создания миров / Пер. с англ. М. В. Лебедева // Гудмен Н. Способы создания миров. М.: Идея-Пресс; Праксис, 2001. С. 253. — Прим. пер.].

соответствующую вещь. Иными словами, релятивист упускает из виду то, что предпосылкой самой мысли является существование некоторого вида объективной «правильности».

Существует интересная связь между аргументом, который я только что подверг анализу (его можно назвать аргументом Платона—Витгенштейна), и аргументом, направленным против тезиса о несоизмеримости, который я приписал Куайну и Дэвидсону. По сути дела, Куайн и Дэвидсон ведут речь о том, что последовательный релятивист не имеет никакого права считать других людей говорящими или мыслящими существами (если издаваемые другими людьми звуки не поддаются пониманию, то они представляют собой звуки и *ничего более*), тогда как Платон и Витгенштейн по сути дела утверждают, что последовательный релятивист не вправе считать себя самого говорящим или мыслящим существом.

ЧТО С ЭТИМ ДЕЛАТЬ?

Аргументы, которые я только что изложил, убедили меня в том, что две наиболее известные разновидности философии науки, созданные в этом столетии, являются противоречивыми. (Конечно, ни одна из них не есть *только* «философия науки»). Это соображение навело меня на мысль порассуждать о значении данной ситуации. Как возникли подобного рода взгляды?

Напомню, что логический позитивизм был одновременно и продолжением предшествовавшего ему позитивизма Маха, и в то же самое время отличался от него. Позитивизм Маха, или «эмпириокритицизм», по большому счету представлял собой попытку выразить эмпиризм Юма при помощи иного словаря. Яркость личности Маха, его догматический и полный энтузиазма стиль и его выдающиеся научные достоинства обусловили широкое культурное влияние его позитивизма. (Ленин, опасаясь того, что большевики обратятся в «эмпириокритицизм», написал книгу, направленную против этого учения). Эйнштейн, чья интерпретация специальной теории относительности была операционалистской по своему духу (в противоположность той интерпретации, которую он дал общей теории относительности), признавал, что его критика понятия одновременности была многим обязана Юму и Маху, хотя, к его разочарованию, Мах

целиком и полностью отверг специальную теорию относительности.

Однако наиболее примечательным событием, которое привело к появлению логического позитивизма, была революция в дедуктивной логике. В 1879 году Фреге открыл алгоритм, механическую процедуру доказательства, которая включала в себя то, что сегодня принято называть логикой второго порядка. Эта процедура является *полной* для элементарной теории дедукции («логика первого порядка»). Тот факт, что можно написать алгоритм доказательства для *всех* общезначимых формул логики первого порядка — алгоритм, который не требует значимого анализа и воспроизводства человеческой психологии — это значимый факт. Он вселял надежду, что можно сделать то же самое и для так называемой индуктивной логики, благодаря чему «научный метод» мог бы оказаться алгоритмом, и эти два алгоритма — алгоритм для дедуктивной логики (который, конечно же, будучи распространен на логику более высоких порядков, оказался *неполным*) и алгоритм для индуктивной логики, который еще надлежало открыть, — могли бы исчерпывающим образом описать или «рационально реконструировать» не только *научную* рациональность, но и всю рациональность, заслуживающую подобного рода названия.

Когда я только начинал свою преподавательскую карьеру в Принстонском университете, мне довелось познакомиться с Рудольфом Карнапом, который в течение двух лет работал в «Institute for Advanced Studies». В один памятный вечер Карнап рассказал мне, как он стал философом. Карнап объяснил мне, что он заканчивал свое обучение, готовясь стать физиком, и изучал логику в семинаре Фреге. Текстом, который изучался в семинаре, были *Principia Mathematica* (представим себе изучение *Principia* Рассела и Уайтхеда под руководством Фреге!). Карнап был очарован символической логикой; не в меньшей степени он был очарован и специальной теорией относительности. В результате он решил сделать своей дипломной работой формализацию специальной теории относительности при помощи системы записи *Principia*. Именно потому, что кафедра физики в Иене не приняла бы подобного рода работы, Карнап, как он мне сказал, и стал философом.

Полученное в настоящее время множество отрицательных

результатов, включая весомые размышления Нельсона Гудмена, свидетельствует о том, что *не может* существовать *полностью* формальной индуктивной логики. Некоторые важные аспекты индуктивной логики могут быть формализованы (хотя адекватность формализации остается спорной), однако всегда сохраняется потребность в суждениях о «разумности», вне зависимости от того, будут ли они построены посредством выбора словаря (или, точнее, посредством *разделения* словаря на «отображаемые» и «не-отображаемые» предикаты) или же каким-то иным образом. В настоящее время никто по сути дела не верит в то, что существует чисто формальный научный метод (об этом см. главу 8).

История, рассказанная мне Карнапом, подтверждает идею, что именно успех, достигнутый в деле формализации дедуктивной логики, сыграл решающую роль. Если этот успех инспирировал подъем логического позитивизма, то не могла ли неудача в формализации индуктивной логики, открытие того, что не существует *алгоритма* для эмпирической науки, повлечь за собой подъем анархизма? Я не буду настаивать на этом предположении; в любом случае, дополнительные факторы, по всей видимости, сыграли свою роль. В то время как Кун все больше смягчал свою точку зрения, Фейерабенд и Мишель Фуко стремились довести ее до крайности. Более того, они имели в виду нечто политическое: как Фейерабенд, так и Фуко связывают наши нынешние институционализированные критерии рациональности с капитализмом, эксплуатацией и даже с сексуальным подавлением. Конечно же, сегодня существует множество причин, благодаря которым крайний релятивизм оказывается привлекательным для людей; идея о том, что все ныне существующие институты и традиции плохи, является одной из них.

Другая причина — это *сциентизм*. Сциентистский характер логического позитивизма совершенно очевиден и лежит на поверхности; однако я полагаю, что и за релятивизмом скрывается определенная форма сциентизма. Теория, согласно которой «рациональность» — это то, что считается таковой в нашей локальной культуре, никогда не принималась целиком и полностью на вооружение мыслителями-«анархистами»; несмотря на это, она представляет собой естественный предел для собственной «анархизму» тенденции. Кроме всего прочего, это ре-

дукционистская теория. Теория, согласно которой рациональность определяется по идеалу компьютерной программы, — это сциентистская теория, инспирированная точными науками; теория, согласно которой рациональность определяется нормами, свойственными каждой отдельной культуре, — это сциентистская теория, вызванная к жизни антропологией.

Я не собираюсь обсуждать здесь надежды, возникшие у некоторых лингвистов — последователей Хомского, будто когнитивная психология способна открыть *внутренний* алгоритм, определяющий рациональность. Лично я полагаю, что это — интеллектуальная мода, которая принесет такие же разочарования, что и надежда логических позитивистов на создание символической индуктивной логики.

Все это наводит на мысль о том, что те проблемы, с которыми сталкивается современная философия, обусловлены отчасти сциентизмом, проникшим в нее в XIX веке, причем проблемы эти затрагивают не одну только философию. Я вовсе не отрицаю важности логики, равно как и значения формальных исследований в области теории подтверждения, или в области семантики естественных языков и т. д. Однако я склонен полагать, что для философии они имеют побочное значение и что пока мы прочно захвачены формализацией, мы будем обречены колебаться между двумя формами описанного мною сциентизма. Обе эти формы сциентизма представляют собой попытки уклониться от задачи вразумительного описания сферы человеческого разума.

6. ФАКТ И ЦЕННОСТЬ

Понятая в достаточно широком смысле, тема факта и ценности является такой темой, которая касается каждого. В этом отношении она резко отличается от многих философских вопросов. Даже самые образованные люди не считают обязательным иметь мнение о том, действительно ли существует реальный мир или же, например, он только кажется таким. Вопросы философии языка, эпистемологии и даже метафизики могут показаться вопросами, которые, хотя и представляют определенный интерес, все же являются необязательными с точки зрения жизни большинства людей. Но вопрос факта и ценности является вопросом принудительного выбора. Любой мыслящий человек *должен* иметь действительное мнение по этому вопросу (который может совпадать или не совпадать с его умозрительным мнением). Если вопрос факта и ценности является вопросом принудительного выбора для мыслящих людей, то один из ответов на этот вопрос — ответ, состоящий в том, что факт и ценность являются полностью разделенными сферами, что дихотомия «фактическое утверждение или ценностное суждение» является абсолютной, — этот ответ принял статус культурного института.

Называя такую дихотомию культурным институтом, я хочу предположить, что фактом является то, что полученный ответ будет по-прежнему оставаться полученным ответом в течение некоторого времени, несмотря на то, что могут сказать об этом философы, и несмотря на то, является ли полученный ответ *правильным* или нет. Даже если я бы смог убедить вас, что разделение факта и ценности не имеет рациональной основы, что эта дихотомия рационально уязвима, или даже если бы какой-то философ лучше меня смог бы доказать это, используя какой-то совершенно убедительный аргумент (разумеется, в философии такого не существует), все же в следующий раз, когда вы вышли бы на улицу, пошли бы на вечеринку с коктейлями или участвовали в обсуждении в каком-либо совещательном органе, членом которого вы можете быть, вы всегда встретите человека, который говорит вам: «Это предлагается в качестве фактического

утверждения или ценностного суждения?» Взгляд, согласно которому вопрос о том, хороши вещи или плохи, лучше они или хуже и т. д. — это не вопрос факта, становится, в некотором смысле, *институционализированным*.

Суть моего аргумента не будет новой. Я собираюсь реабилитировать несколько дискредитированный шаг в споре о факте и ценности, а именно, шаг, который состоит в заявлении, что это различие является по меньшей мере безнадежно размытым, потому что сами фактические заключения и практики научного исследования, на которые мы опираемся в решении, что является фактом, а что нет, предполагают ценности.

Причина того, что этот ход несколько дискредитирован, состоит в том, что существует очевидное возражение. Возражение против мнения, что наука предполагает ценности, является защитной уступкой. Защитники разделения факта и ценности допускают, что наука предполагает некоторые ценности, например, наука предполагает, что нам нужна *истина*, но они утверждают, что эти ценности не являются *этическими* ценностями. Я буду воображать некоторого искусственного противника, который отстаивает мнение, что наука предполагает *одну* ценность, а именно, ценность самой истины.

Как мы видели, истина не является простым понятием. Идея о том, что истина является пассивной копией того, что «наличествует» «реально», т. е. независимо от сознания и от рассуждения, рухнула под ударами критики, инспирированной Кантом, Витгенштейном и другими философами, даже если она по-прежнему глубоко владеет нашим мышлением.

Некоторые философы обратились к *принципу эквивалентности*, т. е. принципу, согласно которому *сказать, что некоторое утверждение истинно, эквивалентно тому, чтобы сделать само это утверждение*, чтобы доказать, что в отношении истины не существует действительных философских проблем. Другие обращались к работе Альфреда Тарского, логика, который показал, как, имея формализованный язык (формальную запись для выражения определенных утверждений, использующей символическую логику) можно определить истину для данного языка в более сильном языке (так называемом «метаязыке»)¹.

1. См. не-технический обзор работы Тарского в моей книге «Значение и моральные науки», ч. 1, лекция 1.

Работа Тарского и сама основывалась на принципе эквивалентности: фактически его критерием успешного определения «истинного» было то, что оно должно давать все предложения формы ««Р» истинно тогда и только тогда, когда Р», например,

(Т) «Снег бел» истинно тогда и только тогда, когда снег бел;

в качестве теорем метаязыка (где Р является предложением рассматриваемой формальной записи).

Но принцип эквивалентности является философски нейтральным, как и работа Тарского. Согласно *любой* теории истины предложение «снег бел» эквивалентно предложению ««снег бел» истинно».

Философы-позитивисты ответили бы, что если вам известно вышеприведенное (Т), то вы знаете, что означает предложение ««снег бел» истинно»: оно означает, что *снег бел*. А если вы не понимаете, что такое «снег» и «бел», добавили бы они, то у вас действительно проблема! Но проблема состоит не в том, что мы не понимаем предложение «снег бел»; проблема в том, что мы не понимаем, *что значит понятие* «снег бел». Это философская проблема. Об этом (Т) не говорит ничего.

И действительно, разве это не соответствует нашим интуитивным представлениям об этой проблеме? Если к нам подойдет некто с блеском в глазах и скажет: «Хотите узнать «Истину»?», то наша обычная реакция состояла бы в том, чтоб огнестись к этому человеку с большим подозрением. И причина нашего подозрения (помимо блеска в его глазах) состоит как раз в том, что кто-то, говорящий нам, что хочет, чтобы мы узнали истину, не говорит нам фактически *ничего* до тех пор, пока у нас нет представления о том, каких стандартов рациональной приемлемости он придерживается: что он считает рациональным способом проведения исследования, каковы его стандарты объективности, когда он считает рациональным прекратить исследование, какие основания он считает достаточными для принятия того или иного заключения по вопросам любого рода, которые могут вызывать у него интерес. В применении к науке я бы сказал, что заявление о том, что наука «стремиться открыть истину», является в действительности чисто формальным утверждением. Оно не говорит ничего, кроме того, что ученые не желают признавать снег белым, если он не белый, что они не

желают утверждать, что электроны текут по проводам, если электроны не текут по проводам, и так далее. Но эти чисто формальные утверждения являются довольно бессодержательными до тех пор, пока у нас нет некоторой идеи, согласно которой именно система критериев рациональной приемлемости отличает научные способы определения белизны снега от других способов определения белизны снега, научные способы определения тока электронов по проводам от других способов определения ток электронов по проводам, и т. д.

Если сравнение нашей системы представлений с неконцептуализованной, беспонятийной реальностью с целью проверить их совпадение не имеет смысла, то требование, что наука стремится открыть истину, может означать только то, что наука стремится построить картину мира, которая, в некотором идеальном пределе, удовлетворяет определенным критериям рациональной приемлемости. То, что наука стремится построить картину мира, которая была бы *истинной*, само по себе является истинным утверждением, почти пустым и формальным истинным утверждением; цели науки приобретают материальное содержание только благодаря критериям рациональной приемлемости, которые используются в науке. Итак, я говорю, что ответ на рассмотренную позицию искусственного спорщика о том, что единственной целью науки является открытие истины (помимо указания на то, что наука имеет и другие цели, что, разумеется, верно), состоит в том, что *истина не является базисным уровнем*: истина сама вырастает из наших критериев рациональной приемлемости, и именно их мы должны рассматривать, если хотим открыть ценности, которые действительно содержатся в науке.

Для примера давайте сейчас вообразим крайний случай разногласия. Разногласие, которое я собираюсь вообразить, не является обычным научным разногласием, хотя я надеюсь, что наш ответ на него позволит открыть нечто в природе научных ценностей.

Гипотеза, по поводу которой в данном случае возникает описываемое разногласие, — это гипотеза, которую мы обсуждали в главе 1, гипотеза о том, что все мы — Мозги в Бочке. Мы доказывали, что эта гипотеза не может быть истинной с некоторой вероятностью; но предположим, что в этом разногласии наши

аргументы не смогли убедить противную сторону (что не является невозможным, так как философские аргументы никогда не убеждают всех). Вкратце, гипотеза состоит в том, что все является коллективной галлюцинацией, описанной нами ранее.

Разумеется, если бы все было нашей коллективной галлюцинацией, то для многих людей это не имело бы никакого значения. Например, не было бы практически никакой разницы для влюбленных². И я представляю, что не было бы никакой разницы для экономистов. (Зачем экономисту беспокоиться о том, все ли деньги в мире являются физически реальными? Большая часть из них не является физически реальной по любой теории!)

Я хочу, чтобы читатель вообразил, будто эта сумасшедшая (и, я утверждаю, несвязная) теория — теория, что все мы — Мозги в Бочке — поддержана не отдельным сумасшедшим, а фактически всеми людьми в некоторой большой стране, скажем, в Австралии. Вообразим, что в Австралии лишь небольшое меньшинство людей верит тому же, что мы, а подавляющее большинство полагает, что мы — Мозги в Бочке. Возможно, австралийцы верят этому, потому что все они — ученики какого-то гуру, скажем, Сиднейского Гуру. Возможно, когда мы заведем с ними речь об этом, они скажут: «О, если бы вы могли поговорить с Сиднейским Гуру, посмотреть в его глаза и увидеть, какой он хороший, добрый, мудрый человек, вы тоже были бы убеждены». И если мы спросим: «Но откуда Сиднейский Гуру знает, что мы — мозги в бочке, если иллюзия так точна, как вы говорите?», то они ответят: «О, уж Сиднейский Гуру *знает*»».

Как я говорил выше, это — не научное разногласие в обычном смысле. Мы можем воображать, что австралийцы не хуже нас предвосхищают опытные данные, не хуже нас строят прочные (или кажущиеся таковыми) мосты и т. д. Они даже могут охотно принимать наши самые последние научные открытия — не как истинные, но как правильные описания того, что кажется происходящим в образе. Мы можем или не можем представить, что они не соглашаются с нами о некоторых предсказаниях относительно самого отдаленного будущего (например, они могли бы ожидать, что однажды автоматические механизмы сло-

2. Однако мое мнение о том, так ли это, не окончательно.

маются, и тогда у людей начнутся коллективные галлюцинации такого вида, который даст неопровержимое свидетельство их правоты)³, но вне зависимости от того, делают ли они такие предсказания или не соглашаются ни с какими предсказаниям, отличными от допустимых в соответствии со стандартной теорией, это не затронет мой аргумент. Дело в том, что здесь я вообразил случай, где большое число людей имеет самостоятельную систему убеждений, которая в своих существенных чертах не согласуется с нашей.

И это не вопрос разногласия в «этических» ценностях; этика австралийцев может как угодно близко походить на нашу. (Хотя древний грек сказал бы, что быть *мудрым* — это *этическая* ценность; иудаизм и христианство фактически сузили понятие этического в силу определенной концепции спасения).

Первое, что я хочу заметить о гипотетических австралийцах, — это то, что их мировоззрение является *сумасшедшим*. Иногда, что и говорить, слово «сумасшедший» употребляется с оттенком одобрения; но я здесь не подразумеваю этот смысл. Я думаю, что мы с большой печалью воспринимали бы сообщество людей со столь безумным мировоззрением. Австралицы были бы расценены как сумасшедшие в смысле обладания *больными* сознаниями; и характеристика их сознаний как больных — этическая, или граничит с этической. Но как, не прибегая к брани, мы можем поспорить с австралийцами? (Или попытаться спорить с ними, поскольку я предполагаю, что они не должны быть убеждены).

Аргумент, который сразу приходит на ум, имеет отношение к *бессвязности* их позиции. Я подразумеваю не только бессвязность, которую мы нашли в представлении в главе 1. То была *глубокая* бессвязность, которая требует выставления философского (а следовательно, спорного) аргумента. Но позиция австралийцев бессвязна на гораздо более поверхностном уровне. Одна из вещей, к которым мы стремимся — это то, что мы должны быть в состоянии объяснить, каким образом мы знаем, что

3. Если они делают такие предсказания, то ситуация меняется: их позиция больше не является противоречивой в том отношении, в каком мы критиковали ее в главе 1, так как они выдвигают требование, которое могло бы быть обосновано (в конечном счете), а следовательно, для его понимания не требуется концепции 'трансцендентной' (или независимой от обоснования) истины.

наши утверждения истинны. Частично мы пробуем делать это, развивая каузальную теорию восприятия, так, чтобы мы могли объяснять то, что мы принимаем за надежность нашего перцептивного знания, рассматриваемого непосредственно изнутри нашей теории, приводя объяснение в пределах теории того, как наше восприятие следует из действия воспринимающих органов под влиянием внешнего мира. Частично мы пробуем делать это в соответствии с теорией статистики и постановки эксперимента, так, чтобы мы могли показать, в пределах самой нашей теории, как процедуры, которым мы следуем, чтобы исключить экспериментальную ошибку, действительно в большинстве случаев имеют тенденцию исключить экспериментальную ошибку. Короче говоря, важное и чрезвычайно полезное требование, предъявляемое непосредственно к самой нашей теории, состоит в том, что развиваемая нами теория мира в целом должна включить теорию той деятельности и тех процессов, при помощи которых мы способны знать, что эта теория правильна.

Система австралийцев, однако, не имеет этого свойства связности (по крайней мере, насколько мы можем судить, причем «связность» — это не нечто, для чего у нас есть алгоритм, а нечто, что мы в конечном счете «нутром чуем»). Австралийцы, как вы помните, самостоятельно постулировали иллюзию столь совершенную, что нет никакого рационального способа, которым Сиднейский Гуру может с некоторой вероятностью *знать*, что система убеждений, которую он принял и убедил принять всех остальных, правильна. Судя по нашим стандартам связности, *их* система убеждений полностью бессвязна.

Можно пополнить список методологических достоинств, которых недостает их системе убеждений. Их система убеждений, как я ее описал, *в образе* согласуется с нашей относительно законов природы; но сообщает ли она нам, действительно ли законы природы, которые, как кажется, действуют в образе — это те законы природы, которые фактически действуют вне бочки? Если она не сообщает нам этого, то ей недостает некоторой всесторонней полноты, к которой мы стремимся, поскольку даже в своих собственных терминах она не сообщает нам, каковы истинные и окончательные законы природы. Конечно, это нарушает принцип бритвы Оккама. К тому же кажется трудным или даже невозможным формализовать принцип бритвы Оккама как

алгоритм, но сам тот факт, что теория Мозгов в Бочке постулирует все виды объектов вне бочки, не играющих никакой роли в объяснении наших опытов, согласно самой теории, проясняет, что это — случай, в котором мы можем определенно сказать, что принцип «не умножать сущности без необходимости» нарушен. Назовем теорию, которая удовлетворяет условиям бритвы Оккама по духу, а не только по букве, *функционально простой*.

Фактически я утверждаю, что процедуры, в соответствии с которыми мы останавливаемся на приемлемости той или иной научной теории, имеют отношение к тому, действительно ли научная теория в целом демонстрирует некоторые «достоинства». Я предполагаю, что процедура построения научной теории не может быть правильно проанализирована как процедура подтверждения научной теории *предложение за предложением*. Я предполагаю, что проверка в науке — холистический вопрос, что есть полные теоретические системы, которые выполняют опытный тест «как единое целое», и что суждение о том, как хорошо полная система предложений проходит полную проверку, — это в конечном счете своего рода интуитивный вопрос, который не может быть формализован, если только не формализована вся человеческая психология. Но вернемся к нашему первоначальному вопросу. Каковы неявные ценности науки?

Я доказывал, что если мы берем ценности, к которым мы обращаемся в нашей критике теоретиков Мозгов в Бочке, и добавляем, конечно, другие ценности, о которых не идет речь в этом случае — например наше желание инструментальной эффективности, которую мы предположительно разделяем с теоретиками Мозгов в Бочке — то мы получаем картину науки, предполагающей богатую систему ценностей. Если мы рассматриваем идеал рациональной приемлемости, который раскрывается при рассмотрении того, какие теории ученые и обычные люди считают рационально приемлемыми, то мы видим, что то, что мы пытаемся делать в науке — это строить представление мира, имеющее характеристики инструментальной эффективности, когерентности, всесторонней полноты и функциональной простоты. Но почему это так?

Я ответил бы, что причина, по которой мы предпочитаем этот вид представления, а не «больной» вид понятийного мира, которым обладают австралийцы, теоретики Мозгов в Бочке,

состоит в том, что наличие системы репрезентаций этого вида есть *часть нашей идеи человеческого преуспевания в познании*, и, следовательно, — часть нашей идеи общего процветания человека, эвдаймонии.

Конечно, если метафизический реализм прав и можно рассматривать цель науки просто как попытку заставить наш понятийный мир «соответствовать» миру самому по себе, то можно утверждать, что когерентность, полнота, функциональная простота и инструментальная эффективность интересуют нас только потому, что они служат установлению этого «соответствия». Но понятие трансцендентного соответствия между нашей репрезентацией и миром самим по себе бессмысленно. Отрицать, что нам нужен этот вид метафизического соответствия с ноуменальным миром, не значит отрицать, что нам нужна обычная эмпирическая пригодность (оцененная в соответствии с нашими критериями рациональной приемлемости) для *эмпирического* мира. Но эмпирический мир, в противоположность ноуменальному миру, зависит от наших критериев рациональной приемлемости (и, конечно, наоборот). Мы используем наши критерии рациональной приемлемости, чтобы создать теоретическую картину «эмпирического мира», а затем, по мере того, как эта картина развивается, мы пересматриваем сами наши критерии рациональной приемлемости в свете этой картины, и так далее до бесконечности. Зависимость наших методов от нашей картины мира — тезис, который я подчеркнул в других моих книгах; здесь же я хочу обратить внимание на другую сторону этой зависимости — зависимость эмпирического мира от наших критериев рациональной приемлемости. Я хочу сказать, что мы должны иметь критерии рациональной приемлемости для того, чтобы вообще иметь эмпирический мир, что они частично раскрывают наше понятие оптимального спекулятивного разума. Короче говоря, я утверждаю, что «реальный мир» зависит от наших ценностей (и, снова, также и наоборот).

ПО КРАЙНЕЙ МЕРЕ НЕКОТОРЫЕ ЦЕННОСТИ ДОЛЖНЫ БЫТЬ ОБЪЕКТИВНЫ

Тот факт, что наука не является «нейтральной в отношении ценностей», как это принято считать, конечно, не показывает, что

«этические» ценности объективны или что этика может быть наукой. Фактически, нет никакой перспективы «науки» этики, будь то в смысле лабораторной или дедуктивной науки. Аристотель уже давно отметил, что:

«При подобных предметах рассуждений и подобных предпосылках желательно приблизительно и в общих чертах указать на истину, а если рассуждают о том, что имеет место лишь в большинстве случаев и при соответствующих предпосылках, то [довольно уже и того, чтобы] и выводы [распространялись лишь на большинство случаев]. Конечно, таким же образом нужно воспринимать и каждое наше отдельное высказывание; ведь человеку образованному свойственно добиваться точности для каждого рода [предметов] в той степени, в какой это допускает природа предмета. Одинаково [нелепым] кажется как довольствоваться правдоподобными рассуждениями математика, так и требовать от ратора строгих доказательств»⁴.

Но факт, что рациональная приемлемость в точных науках (являющихся, конечно, центральными примерами рационального размышления) зависит от таких познавательных *достоинств* как «когерентность» и «функциональная простота», показывает, что по крайней мере некоторые ценностные термины относятся к свойствам вещей, к которым они приложимы, а не только к чувствам человека, использующего эти термины.

Если термины «когерентный» и «простой» относятся не к *свойствам* теорий, даже нечетких или несовершенно определенных, но только к «позициям» некоторых людей по отношению к теориям, то такие термины для рациональной приемлемости как «обоснованный», «хорошо подтвержденный», «лучшее из доступных объяснений» должны также быть полностью субъективны, поскольку рациональная приемлемость не может быть более объективна, чем параметры, от которых она зависит. Но, как мы доказывали в предыдущей главе, идея, что сама рациональная приемлемость просто субъективна, опровергает себя. Поэтому мы вынуждены заключить, что по крайней мере *эти* ценностные термины имеют некоторое объективное применение, некоторый вид объективных условий обоснования.

Конечно, можно попытаться избежать признания объектив-

4. Ethica Nicomachea, кн. 1, гл. 3.

ных ценностей любого вида, отрицая, что «когерентный», «простой», «обоснованный» и т. п. — термины *ценностей*. Можно считать, что они относятся к свойствам, которые мы оцениваем, но что нет никакой объективной правильности нашей оценки. Но эта линия сразу сталкивается с трудностями. «Когерентный» и «простой» имеют слишком много общих характеристик с парадигматическими ценностными словами. Подобно «доброму», «красивому» и «хорошему», «когерентный» и «простой» часто используются как термины *похвалы*. Наши концепции когерентности, простоты и обоснования так же исторически обусловлены, как и наши концепции доброты, красоты и совершенства; эти эпистемические термины фигурируют в тех же видах постоянных философских споров, что и термины этических и эстетических ценностей. Концепция рациональности кардинала Джона Ньюмена очевидно весьма отлична от концепции рациональности Рудольфа Карнапа. Крайне маловероятно, чтобы один из них мог убедить другого, даже если бы они жили в одно то же самое время и могли встретиться. Вопрос, *какая из них является рациональной концепцией самой рациональности*, труден *только* в том же отношении, в каком является трудным обоснование этической системы. Нет никакой *нейтральной* концепции рациональности, к которой мы могли бы обратиться.

Здесь можно попробовать различные конвенционалистские ходы, например, назвать «обосновано _{по Карнапу}» одним «свойством», а «обосновано _{по Ньюмену}» — другим, и что «субъективное ценностное суждение» участвует в решении означать «обосновано _{по Карнапу}» или «обосновано _{по Ньюмену}» в силу использования слова «обосновано», но что никакое ценностное суждение не участвует в утверждении того факта, что данное утверждение *обосновано по Карнапу* или *обосновано по Ньюмену*. Но с чьей точки зрения используется слово «факт»? Если нет никакой концепции рациональности, которую объективно *должны* разделять все, то понятие «факта» лишено содержания. Без когнитивных ценностей когерентности, простоты и инструментальной эффективности у нас нет никакого мира и никаких «фактов», даже фактов о том, что является *относительным* к чему, поскольку они находятся в одной лодке со всеми прочими фактами. И было бы произволом считать эти когнитивные ценности не принадлежащими холистической концепции человеческого процветания.

Когда мы лишены старой реалистской идеи истины как «корреспонденции» и позитивистской идеи обоснования, проводимого в соответствии с общественными «критериями», мы остаемся перед необходимостью рассматривать наш поиск лучших концепций рациональности как целенаправленную человеческую деятельность, которая, подобно всякой деятельности, понимающейся над привычкой и простым следованием склонности или навязчивой идее, руководствуется нашей идеей блага.

РАЦИОНАЛЬНОСТЬ В ДРУГИХ ОБЛАСТЯХ

Если ценности, неявные в науке, особенно в точных науках, частично показывают нашу идею блага, то я думаю, что остальная часть нашей идеи блага может быть выражена нашими стандартами рациональной приемлемости в других областях знания. В этом пункте, однако, необходимо расширить понятие *стандартов рациональной приемлемости*.

До сих пор мы рассматривали лишь стандарты рациональной приемлемости в буквальном смысле: стандарты, которые сообщают нам, когда мы должны и когда не должны *принять* те или иные утверждения. Но стандарты рациональности в широком смысле должны касаться не только того, как мы оцениваем истинность или ложность систем утверждений, но также и того, как мы судим их *адекватность* и *отчетливость*. Есть способы — вполне когнитивные — которыми система утверждений может оказаться не в состоянии предоставить нам удовлетворительное описание, кроме тех случаев, когда она ложна.

Если бы у меня был выбор, я мог бы сделать это замечание даже в связи с теоретической наукой. Я мог бы заметить, что задача точной науки состоит не в том, чтобы обнаружить утверждения, которые являются истинными, или даже утверждения, которые являются истинными и универсальными по форме («законы»), но в том, чтобы найти утверждения, которые являются истинными и *уместными*. А понятие *уместности* влечет за собой широкий набор интересов и ценностей. Но это лишь доказывало бы, что наше *знание* мира предполагает ценности, а не выдвигало бы более радикальное требование, что то, что *считается* реальным миром, зависит от наших ценностей.

Когда мы доходим до перцептивной рациональности, т. е. до

неявных стандартов и навыков, на основе которых мы решаем, может ли кто-то дать истинную, адекватную и ясную теорию даже самых простых *перцептивных* фактов, то мы видим, что здесь действует большое количество факторов. Недавно физиологи подчеркнули, сколько теоретического конструирования вовлечено в даже самые простые случаи восприятия. Мало того, что это истинно на нейрофизиологическом уровне, но и это также истинно и на уровне культуры. Представитель культуры, не знающей мебели, может войти в комнату и дать некоторое ее описание, но если бы он не знал, что из себя представляет стол, стул или парта, то его описание едва ли выразило бы какую-то информацию, которую другой представитель этой культуры хотел бы получить об этой комнате. Его описание состояло бы из истинных утверждений, но было бы неадекватным.

Этот простой пример показывает, что требование адекватности описания неявно представляет собой требование, чтобы описывающий имел в своем распоряжении определенное множество *понятий*; мы ожидаем от рациональных авторов, в отношении к определенным видам *описаний*, что они смогут приобрести определенные понятия и смогут оперировать ими. Тот факт, что описывающий не употребил определенное понятие, может дать основание для критики как его самого, так и проведенного им описания.

То, что является истинным на простом уровне разговора о столах и стульях в комнате без людей, истинно также и на уровне описания межличностных отношений и ситуаций. Рассмотрим термины, которые мы используем каждый день в описании того, каковы другие люди, например, *сдержанные* или *несдержанные*. *Сдержанный* и *несдержанный* могут, конечно, использоваться для похвалы или обвинения; и одно из многих различий, которые перепутались под общим заголовком «различие ценности и факта», является различием между употреблением языкового выражения для описания и употреблением языкового выражения для похвалы или обвинения. Но это — не то различие, которое может быть проведено на основе словаря. Суждение, заключающее, что кто-то невнимателен, может действительно использоваться для обвинения; но оно может использоваться просто для описания, а также может использоваться для объяснения или предсказания.

Например, я могу сказать вам: «Не позволяйте Джонсу задевать ваши чувства. Вы, вероятно, думаете, что он невзлюбил вас, судя по тому, как он с вами разговаривает, но это — обычное неверное впечатление. Независимо от того, как он к вам относится, он будет скорее всего вести себя таким образом, что ваши чувства будут задеты. Он просто довольно несдержанный человек, но не думайте, что это имеет какое-либо отношение к вам лично».

В этой небольшой мнимой речи некто использует слово «несдержанный» не с целью обвинить Джонса, а с намерением предсказания и объяснения поведения Джонса по отношению к кому-то еще. Как предсказание, так и объяснение могут быть совершенно *правильны*. И точно так же «ревнивый» может быть термином вины и может использоваться вообще без намерения обвинения. (Иногда можно иметь полное право быть ревнивым.)

Использование слова «несдержанный» кажется мне прекрасным примером того отношения, в котором различие факта/ценности в реальном мире и в реальном языке является безнадежно нечетким. Важность таких терминов, как «несдержанный», «дерзкий», «упрямый», «надоедливый», и т. д. в фактической моральной оценке была подчеркнута Айрис Мердок в книге «Суверенитет «Добра»»⁵. Даже при том, что каждое из утверждений «Джон — очень несдержанный человек», «Джон не думает ни о ком, кроме себя», «Джон сделал бы за деньги практически что угодно» может быть просто истинным описанием в наиболее позитивистском смысле (причем заметьте, что «Джон сделал бы за деньги практически что угодно» не содержит никаких ценностных терминов), если мы утверждаем конъюнкцию этих трех утверждений, то едва ли необходимо добавлять «Джон — не слишком хороший человек». Когда мы думаем о фактах и ценностях как о не зависящих друг от друга, мы обычно думаем о «фактах» как о выраженных на некотором физикалистском или бюрократическом жаргоне, а о «ценностях» — как о выраженных в самых абстрактных ценностных терминах, например, «хороший», «плохой». Независимость ценности от факта тяжелее поддержать, когда это факты порядка «несдержанный», «думает только о себе», «сделал бы что угодно за деньги».

5. Murdoch A. The Sovereignty of Good. Routledge and Kegan Paul, 1970.

Подобно тому, как мы критикуем автора описания, не использующего понятия *стола* и *стула*, когда их использование требуется, так же кто-то, кто будет не в состоянии заметить, что кто-то *внимателен* или *непосредственен*, может оказаться уязвимым для критики, состоящей в том, что он ненаблюдателен или поверхностен, а его описание не адекватно.

СУПЕР-БЕНТАМИТЫ

Позвольте мне вернуться к моему прошлому примеру со сторонниками теории о Мозгах в Бочке и изменить его. На сей раз представим себе, что австралийский континент населен людьми такой культуры, которая разделяет нашу историю, географию и точные науки, но которая не соглашается с нами в этике. Я не хочу брать обычный случай супернацистов или чего-либо в этом роде, а возьму значительно более интересный случай супер-Бентамитов. Представим себе, что австралийский континент населен людьми, разработавшими некоторую научную меру того, что они принимают за «гедонистический стиль», и они полагают, что нужно всегда действовать так, чтобы максимизировать гедонистический стиль (под этим подразумевается самый большой гедонистический стиль самого большого числа людей). Я предположу, что супер-Бентамиты чрезвычайно искушены, знают все трудности предсказания будущего и точной оценки последствий действий и т. д. Я также предположу, что они чрезвычайно безжалостны, и что, в то время как они не причинили бы кому-либо страданий ради самого большого счастья самого большого числа людей, если бы было разумное сомнение, что *фактически* последствия их действий *должны* были бы вызвать самое большое счастье самого большого числа людей, но в тех случаях, когда им точно известно, каковы будут последствия действий, они легко совершают самые ужасные действия — будут мучить маленьких детей или осуждать людей за преступления, которых те не совершали, — если результат этих действий должен будет в конечном счете увеличить общий уровень удовлетворения (после внесения в каждом случае должной поправки на страдание невинной жертвы) на любое положительное ϵ , хотя бы и небольшое.

Я думаю, что мы были бы не в восторге от супер-Бентамитс-

кой морали такого вида. Большинство из нас осудило бы супер-Бентамитов как пользующихся порочной системой ценностей, как бюрократов, как безжалостных людей и т. д. Они — «новый человек» в его наиболее ужасном проявлении. А они отплатили бы нам подобными же обличениями, заявляя, что мы глуповаты, суеверны, пленники иррациональной традиции и т. д.

Разногласие между нами и супер-Бентамитами — это разногласие такого вида, которое обычно предполагается, чтобы заметить, что две группы людей могли бы договориться обо всех фактах и все еще не соглашаться о «ценностях». Но рассмотрим этот случай подробнее. Каждый супер-Бентамит знаком с фактом, что иногда самое большое удовлетворение самого большого числа людей (измеренное в «утилях») требует, чтобы кто-то солгал. И сказать неправду в целях максимизирования общего уровня удовольствия не считается «нечестностью» в бранном смысле. Так что некоторое время спустя использование слова «честный» среди супер-Бентамитов чрезвычайно отличалось бы от использования этого же самого описательного термина среди нас. И то же самое произойдет с терминами «сдержанный», «хороший гражданин» и т. д. Словарь, доступный супер-Бентамитами для описания ситуаций человеческих взаимоотношений, будет совершенно отличен от словаря, доступного нам. Многие из наших описательных средств будут у них отсутствовать или будут изменены до неузнаваемости, но, мало того, они еще и изобретут, скорее всего, свой собственный новый жаргон (например, точные термины для описания гедонистического стиля), недоступный для нас. Фактура человеческого мира начнет изменяться. Кончится это тем, что со временем супер-Бентамиты и мы окажемся живущими в разных человеческих мирах.

Короче говоря, дело не в том, что супер-Бентамиты и мы «согласны по поводу фактов и не согласны по поводу ценностей». Почти во всех межличностных ситуациях наши описания фактов будут радикально отличаться от их описаний фактов. Даже если ни одно из тех утверждений, которые они делают о некоторой ситуации, не *ложно*, их описание будет не таким, какое мы сочли бы адекватным и ясным; а то описание, которое дадим мы, будет не таким, какое сочли бы адекватным и ясным они. Короче говоря, даже если мы отвлечемся от наших «раз-

погласий по поводу ценностей», мы не сможем рассматривать их общую репрезентацию человеческого мира как полностью рационально приемлемую. И так же, как неспособность сторонников теории Мозгов-в-Бочке правильно понять, *каким образом существует мир*, является прямым результатом их порочных стандартов рациональности — порочных стандартов теоретической рациональности, так и неспособность супер-Бентамитов правильно понять, *каким образом существует человеческий мир*, является прямым результатом их порочной концепции человеческого процветания.

СУБЪЕКТИВИСТСКАЯ ТЕОРИЯ ДОБРОДЕТЕЛИ

Часто считается, что шаг от утверждения «Джон внимателен, правдив, добр, храбр, ответственен и т. д.» к утверждению «Джон морально добродетелен» включает по крайней мере одну недоказанную (и недоказуемую) «посылку», а именно: «Сдержанность представляет собой моральную добродетель». Также считается, что необходимость в моральных «посылках» перед тем, как выводить моральные заключения из «фактуальных» утверждений, показывает, что этические суждения не могут быть рационально обоснованы.

Такое изображение этики в виде перевернутой пирамиды с вершиной (в свою очередь, ни на чем не покоящейся) из «этических аксиом», поддерживающей весь наш корпус моральных убеждений и размышлений, наивно. Никому еще не удавалось наложить на этику аксиоматическую структуру (как заметил Аристотель в цитированном мною выше фрагменте, подобные моральные правила почти всегда истинны лишь «в общих чертах»). Ту же самую уловку — изображение корпуса размышления, которое хотят поставить под сомнение, покоящимся на неприемлемых «аксиомах» — скептики использовали в каждой области. Например, скептики, сомневающиеся в существовании материальных объектов, доказывают, что принцип, состоящий в том, что «если наши ощущения возникают так, как они возникали бы, если бы существовал материальный мир, то, по всей вероятности, существует материальный мир», является *рационально неприемлемой посылкой*, которую мы неявно привлекаем всякий раз, когда утверждаем, что «наблюдаем» материальный объект, или

пытаемся иным образом обосновать веру в их существование. Фактически, этика, как и математика, и рассуждение о материальных объектах предполагает понятия, а не «аксиомы». Понятия используются в наблюдении и обобщении и самостоятельно легитимированы тем успехом, с которым мы используем их для описаний и выводов.

Более продуманная критика идеи этической объективности допускает, что наши этические убеждения опираются на наблюдения частных случаев, «интуитивные представления», общие принципы и т. д., а не на некоторое собрание произвольных «этических аксиом», но выдвигает обвинение в том, что этическое «наблюдение» само инфицировано неизлечимой болезнью: *проекцией*.

Согласно этой теории, люди от природы сострадательны, хотя и не всегда. Так, когда мы видим какой-то страшный случай, скажем, кто-то пытается маленького ребенка для собственно садистского удовольствия, то нас это ужасает. Но психологический механизм «проекции» приводит нас к тому, чтобы испытывать качество чувства как качество самого деяния: мы говорим «поступок был ужасен», когда мы должны были бы в действительности сказать «моя реакция состояла в том, что я ужаснулся». Таким образом мы строим корпус того, что мы принимаем за «этические наблюдения», которые в действительности являются лишь наблюдениями наших собственных субъективных этических чувств.

Эта история (как и любая другая) существует и в более сложной форме. Юм постулировал наличие у людей тенденции, которую он назвал «симпатией» и которая постепенно усиливается под влиянием культуры. Современные социобиологи постулируют наличие у людей инстинкта, который они называют «альтруизмом», и говорят об «альтруистических генах». Но ключевая идея остается прежней: есть этические чувства, но нет никаких объективных ценностных свойств.

Мы уже видели, что это не так: есть по крайней мере *некоторые* объективные ценности, например, *обоснование*. Кроме того, можно утверждать, что *этические* ценности являются субъективными, в то время как *познавательные* ценности объективны; но аргумент, заключающий, что вообще не может быть никаких объективных ценностей, не выдерживает критики.

Чтобы показать недостатки аргументации в пользу морально-го субъективизма, я должен теперь вспомнить аргументы, которые использовались против метафизического реализма в главе 2. Это может показаться странным: разве субъективизм — не *противоположность* метафизического реализма? Если мы так считаем, то нам будет казаться, что любой аргумент против метафизического реализма должен *поддерживать* субъективизм; стратегия, которой я собираюсь следовать, используя *тот же самый* аргумент против и метафизического реализма, и субъективизма, покажется невозможной.

Но, фактически, метафизический реализм и субъективизм — не просто «противоположности». Сегодня мы склонны к чрезмерному реализму в физике и чрезмерному субъективизму в этике, и эти тенденции связаны друг с другом. Именно *потому*, что мы склонны к чрезмерному реализму в физике, потому, что мы рассматриваем физику (или некоторую гипотетическую физику будущего) как Единую Истинную Теорию, а не просто как рационально приемлемое описание, пригодное для некоторых задач и целей — именно поэтому мы склонны к субъективизму в описаниях, которые мы не можем «свести» к физике. Меньшая склонность к реализму в физике и меньшая склонность к субъективизму в этике связаны аналогичным образом.

Аргумент в конце главы 2 был направлен против «физикалистской» или натуралистической версии метафизического реализма. Чтобы напомнить его, предположим, что стандартная интерпретация I (согласно которой «кошка» относится к кошкам, «вишня» к вишням и т. д.) либо коэкстенсивна, либо идентична физикалистскому отношению R . Так, R наличествует между токенами «кошка» (или физическими событиями уместного использования кем-либо этих токенов) и кошками и т. д. Ненормативная интерпретация J , которую мы описали, будет тогда также коэкстенсивна с некоторым отношением R' , определенным в терминах R и возможных миров и перестановок, используемых при построении J (см. Приложение). Так, R' наличествует между токенами «кошка» (или физическими событиями использования кем-либо этих токенов стандартным способом) и *вишнями*, и т. д. Как R , так и R' являются отношениями «корреспонденции» (соответствия): одни и те же предложения «истинны» согласно обоим соответствиям. Действия, которых требу-

ет R' -истинность предложения (то есть действия, которые будут «успешны», с точки зрения агента) — это те же действия, которые требуются, когда предложение R -истинно. Если R «идентично с референцией»; если R , R' , и все прочие отношения, назначающие нашим словам экстенсионалы такими способами, которые удовлетворяют нашим операциональным и теоретическим ограничениям, не являются в одинаковой степени корректными; если R , R' и другие отношения не являются в одинаковой степени корректными потому, что одно из них — R — как раз и является референцией, — то сам этот факт необъясним с физикалистской точки зрения.

Этот аргумент — не просто аргумент против (физикалистской версии) метафизического реализма, но аргумент против *редукционизма*. Если в физикалистской картине мира ничто не соответствует тому очевидному факту, что «кошка» указывает на кошек, а не на вишни, то это — решающая причина для того, чтобы отклонить требование, согласно которому все используемые нами понятия были редуцированы до физических терминов. Ведь референция и истина — понятия, от которых мы не можем последовательно отказаться. Если я думаю: «Кошка лежит на коврике», то я полагаю, что «кошка» на что-то указывает (хотя при этом я не обязательно придерживаюсь теории «референции» в духе метафизического реализма) и полагаю, что «Кошка лежит на коврике» истинно (хотя и не обязательно при этом придерживаюсь теории истины в духе метафизического реализма).

Вернувшись таким образом к аргументу из второй главы, рассмотрим теперь, как он опирается на аргументы в пользу морального субъективизма. Теория «отображения» дает такое объяснение морального опыта: моральный опыт представляет собой, если можно так выразиться, неправильно размещенное субъективное чувство. Сравним по контрасту теорию «отображения» с такой теорией: «до определенной степени у всех людей есть чувство справедливости и некоторая идея блага. Мы отвечаем (периодически) на такие призывы как «будьте добры к странникам, ибо вы знаете, каково быть странником в земле Египетской»⁶. Наша симпатия становится полнее отчасти в силу нашего убеж-

6. Парафраз библейского стиха — Исход 22:21. — *Прим. пер.*

дения в том, что она *должна* быть более полной; мы чувствуем, что жестокость (иногда) *неправильна*, даже когда жертва — такой человек, которому мы с трудом можем искренне сочувствовать. Мы обнаруживаем подобия между увечьями других и нашими собственными увечьями, между чужой выгодой и нашей собственной. Мы изобретаем моральные термины для нравственно важных особенностей ситуаций и постепенно начинаем делать явные моральные обобщения, которые ведут к дальнейшему уточнению наших моральных понятий, и так далее.

Эта теория на первый взгляд, проще и изощреннее, чем теория «отображения». (В частности, она подтверждает роль *аргументации* в формировании моральных позиций.) Однако, многие разумные люди придерживаются того мнения, что в наши дни следует отклонить постановку вопроса о «чувстве справедливости» и об «идее блага» (там, где это не имеет вполне субъективный смысл) как «ненаучную». Итак, моральное знание становится проблематичным, а то и прямо невозможным.

Но что значит здесь «ненаучный»? Убеждение в том, что есть такая вещь, как справедливость — не вера в *духов*, а «чувство справедливости» — не паранормальная способность, которая позволяет нам чувствовать таких духов. Справедливость не является чем-то, что можно предложить *добавить* к списку объектов, допускаемых физикой, как химики XVIII века предлагали добавить «флогистон» к списку объектов, допускаемых в соответствии с химической теорией. Этика не находится в *противоречии* с физикой, как предполагает термин «ненаучный»; дело просто в том, что «справедливый», «благо» и «чувство справедливости» являются понятиями в дискурсе, *не сводимом* к физическому. Как мы только что убедились, *другие* виды сущностного дискурса не сводимы к физическому дискурсу, но не являются незаконными по этой причине. Рассуждения о «справедливости», равно как и о «референции», могут быть *вне*научным, не будучи при этом *ненаучными*.

Как еще один способ увидеть, что тут происходит, рассмотрим любой основной принцип логики или математики, скажем, принцип, утверждающий, что ряд целых чисел всегда может быть продолжен («за каждым числом следует другое число»), или принцип, утверждающий, что непустое множество целых чисел должно содержать наименьший член. Предположим, что

кто-то выдвигает следующее представление: «Эти принципы истинны для чисел и множеств чисел, с которыми мы имеем дело на практике. Поэтому они и кажутся необходимыми. С помощью механизма, который называется «отображением», мы прилагаем это *ощущение необходимости* к самим принципам; мы чувствуем, что *утверждения* имеют таинственное свойство «необходимости». Но в действительности это никак не обосновано. Насколько мы можем судить, эти принципы даже не могут быть истинны».

В сущности, никто не согласился бы с этим. Фактически каждый математик сказал бы вместо этого что-нибудь вроде: «Большинство людей в той или иной степени обладает математической интуицией. Поэтому мы интуитивно «видим» или можем «увидеть» из примеров (или с помощью искусного вопрошания, как мальчик-раб в диалоге Платона), что принципы необходимо истинны».

Курт Гёдель полагал, что «математическая интуиция» аналогична *восприятию*: математические объекты (которые он называл «понятиями») находятся *вне* нас, и наша интуиция позволяет нам интеллигибельно воспринимать эти сущности Платона; но немногие математики согласились бы с такой платонистской метафизикой. Сравнение Гёделем математической интуиции с восприятием исходит из слишком упрощенной идеи восприятия. В действительности зрение дает нам не прямой доступ к готовому миру, а описание объектов, которые частично структурированы и конституированы самим зрением. Если мы примем радугу в представлении физика за радугу «в себе», то окажется, что у радуги «в себе» нет никаких *полосок* (спектроскопический анализ дает гладкое распределение частот); красные, оранжевые, желтые, зеленые, синие и фиолетовые полосы — особенность *перцептивной* радуги, а не радуги физика. Перцептивная радуга зависит непосредственно от природы нашего перцептивного аппарата, от нашего визуального «создания мира», как назвал это Нельсон Гудмен⁷. («Объекты» физика также зависят от нашего создания мира, как показывает изобилие радикально различных версий, которые физики строят об «одних и

7. См.: Гудмен Н. Способы создания миров / Пер. с англ.. М. В. Лебедева. М.: Идея-Пресс, Праксис, 2001. — *Прим. пер.*

тех же» объектах.) Все же мы не считаем зрение *дефектным* потому, что оно видит в радуге полосы; наоборот, дефектное зрение у того, кто не видит их. Зрение делает хорошим его способность дать описание, которое соответствует объектам *для нас*, а не метафизическим вещам-в-себе. Зрение хорошо, когда оно позволяет нам видеть мир «как он есть» — то есть человеческий, функциональный мир, который частично создан самим зрением.

Предложенная новая аксиома теории множеств, типа «Аксиомы выбора», может быть принята частично из-за ее согласия с «интуитивными представлениями» опытных математиков и частично за ее результативность. Если аксиома выбора не давала бы результатов, которые математики сочтут успешными, тот факт, что некоторые люди находят ее «интуитивной», не имел бы большого значения. Сама математическая интуиция демонстрируется или проверяется, схватывая математические принципы и следуя доказательствам. Короче говоря, математическая интуиция хороша, когда позволяет нам видеть математические факты такими, «каковы они есть» — то есть такими, каковы они есть в математическом мире, созданном человеческой математической практикой (включая применение математики к другим предметам).

Физиологическое или психологическое описание зрения не может сообщить нам, можно ли считать наблюдение полос в радуге «правильным» наблюдением или нет, и в еще меньшей степени может физиологическое или психологическое описание мозгового процесса, который происходит, когда кто-либо «схватывает» Принцип Математической Индукции, сообщить нам является ли этот принцип истинным или нет. Как только мы поняли это, нас более не должно удивлять, что описание мозгового процесса, происходящего, когда кто-либо «видит», что то или иное действие является несправедливым, не может сообщить нам, является ли это действие действительно несправедливым.

Разговор о моральном «восприятии», как и разговор о математической интуиции или о референции и понимании, не сводим к языку или картине мира физики. Это не подразумевает, что физика «неполна». Физика может быть «полна» — то есть полна для физических целей. Законченность, которой физике

недостает, — это та же законченность, которой недостает всем конкретным теориям, картинам и дискурсам, так как ни одна теория или картина не полна для всех целей. Если несводимость этики к физике показывает, что ценности суть отображения, то *цвета* — тоже отображения, как и натуральные числа; как и, в этом отношении, «физический мир». Но быть отображением в *этом* смысле — не то же самое, что быть субъективным⁸.

АВТОРИТАРИЗМ И ПЛЮРАЛИЗМ

Я доказывал, что нам необходимы стандарты рациональной приемлемости для того, чтобы у нас вообще был какой-либо мир, будь то мир «эмпирических фактов» или мир «ценностных фактов» (мир, в котором есть красота и трагедия). Должно быть совершенно ясно, что невозможно одновременно и иметь стандарты рациональной приемлемости, и не принимать их или дистанцироваться от них. (Вид скептицизма, который состоит в отказе от любых стандартов рациональной приемлемости, приводит к отказу от любых понятий вообще. Как показал Секст Эмпирик, такой вид эмпиризма в конечном счете невыразим в языке.) Мы имеем такое же право считать нездоровым некоторый «оценочный» склад ума (и все мы так и поступаем), как и считать нездоровым некоторый «познавательный» склад ума. Но это не еще ведет к тому, чтобы отклонить плюрализм или соглашаться с авторитаризмом.

Даже в науке, исходя из того, что наука — объективное предприятие (по стандарту «объективности», который является по общему признанию антропоцентричным, но, как однажды заметил Дэвид Уиггинс, это единственный доступный нам стандарт

8. Неумышленно забавную версию теории отображения приводит К. С. Льюис в «Освобождении Человека» (Macmillan, 1947). Льюис цитирует школьный учебник английского языка (который он из милосердия не называет). «Вспомните двух туристов [*Льюис говорит об известной истории про Кольриджа и водопад*]: один назвал это «возвышенным», а другой «симпатичным»; Кольридж мысленно подтвердил первое суждение и отклонил второе с отвращением. Лай и Тит [*псевдонимы, которые Льюис дал авторам текста*] комментируют следующим образом: «Когда человек сказал, что это возвышенно, он, казалось, говорил о водопаде... Фактически он говорил не о водопаде, но о своих собственных чувствах. В действительности он говорил вот что: *В моем сознании есть чувства, связанные со словом «возвышенное», или, короче, «У меня возвышенные чувства»*».

объективности), мы еще не должны считать, что на каждый научный вопрос есть определенный ответ. На некоторые научные вопросы могут существовать *объективно неопределенные* ответы, то есть не может быть никакой сходимости относительно ответа на них даже в идеальном пределе научного запроса; а некоторые научные вопросы могут иметь определенные, но относительные к контексту ответы (например, на вопрос «Что было причиной сердечного приступа Джона?» можно дать различные правильные ответы в зависимости от того, кто задает вопрос и почему). И, точно так же, считать, что этический запрос объективен в том смысле, что некоторые «ценностные суждения» являются определенно истинными, а некоторые определенно ложны, и, шире, что некоторые ценностные позиции (и некоторые «идеологии») определенно неправильны, а некоторые — определенно ниже некоторых других, — это не то же самое, что разделять глупое мнение, что нет вообще никаких неопределенных случаев. (Один особенно важный вид неопределенного случая был подчеркнут Бернардом Уильямсом: это такой случай, где все альтернативы настолько ужасны, что нет ни одной альтернативы, которая могла бы быть четко выбрана идеально рациональным и мудрым человеком). А зависимость от контекста в этике разумеется сама собой.

Если сегодня мы отличаемся от Аристотеля, то это отличие состоит в том, что мы гораздо более плюралистичны, чем был Аристотель. Аристотель признавал, что различные идеи эвдаимонии, различные концепции процветания человека могут подходить различным индивидуумам из-за различия в их организации. Но, как представляется, он думал, что есть некоторый идеальный вид организации, которую должен иметь каждый человек; что в идеальном мире (опуская мирской вопрос о том, кто выращивал бы зерно и кто пек бы хлеб) каждый человек был бы философом. Мы соглашаемся с Аристотелем в том, что различные идеи процветания человека подходят индивидуумам с различными организациями, но мы идем далее и полагаем, что даже в идеальном мире присутствовали бы различные организации, что разнообразие является частью идеала. И мы видим между идеалами некоторую трагическую напряженность, видим, что следование некоторым идеалам всегда исключает следование некоторым другим. Но, подчеркнем снова, вера в плюрали-

стический идеал — не то же самое, что вера в то, что каждый идеал процветания человека столь же хорош, как и всякий другой. Мы можем отклонять те или иные идеалы процветания как неправильные, как инфантильные, как нездоровые, как односторонние.

Так же не следует путать и обязательство к этической объективности с совершенно иным вопросом — обязательством к этическому или моральному авторитаризму. Возможно, именно это смешение привело одного выдающегося философа⁹ к тому, чтобы поддержать позицию, которую сам он считает ограниченной версией «не-когнитивизма», и утверждать, что «в вопросе о том, что такое «жить наиболее полно», для каждого человека конечной инстанцией должен быть сам человек». (Отметим двусмысленность используемого здесь выражения «конечная инстанция»: подразумевает ли оно конечную *политическую* инстанцию? Конечную *эпистемологическую* инстанцию? Или же оно подразумевает, что *здесь нет никакого фактического вопроса*, как это предлагает его использование термина «не-когнитивизм»?) Уважение к личностям как автономным моральным агентам требует, чтобы мы предоставили им право выбирать моральную точку зрения для себя, как ни отвратителен для нас мог бы быть их выбор. Согласно философии политического либерализма, оно также требует от нас и настояния на том, чтобы правительство не резервировало за собой приоритет на индивидуальный моральный выбор, устанавливая государственную религию или государственную этику. Но радикальная оппозиция ко всем формам политического и морального авторитаризма не требует от ее сторонников морального релятивизма или морального скептицизма. Причина того, что правительство не должно диктовать этику индивидууму, состоит не в том, что не может быть никакого фактического вопроса о том, какие формы жизни полны, а какие формы жизни неполны или нравственно неправильны в некотором другом отношении. (Если бы мораль не могла быть неправильной, то для правительства не было бы *неправильным* регулировать моральный выбор). Многие боятся, что если они признают вслух любой вид моральной объективности, то неко-

9. Wiggins D. Truth, Invention, and the Meaning of Life // Proceedings of the British Academy. Vol. LXII, 1976.

торое правительство может заставить их принудительно принять *свое* понятие моральной объективности — и в этом, безусловно, одна из причин того, почему так много людей признают моральный субъективизм, с которым они не обнаруживают никакого согласия в реальности.

7. РАЗУМ И ИСТОРИЯ

Одновременно с успехами науки пришло понимание того, что многие проблемы не поддаются разрешению при помощи методов точных наук. Наиболее показательными примерами этого являются идеологические и этические проблемы. Поэтому наряду с ростом нашего восхищения и уважения к физическому, космологу, молекулярному биологу росло недоверие к политическому мыслителю, моралисту, экономисту, музыковеду, психиатру и т. д.

В этой ситуации многие интеллектуалы решили идти в ногу со временем и, починяясь культурной моде своих дней, принялись утверждать, что только точные науки дают нам знание. Соответственно, предполагалось, что социальные науки дают знание лишь в той степени, в какой они подражают точным наукам. Эта точка зрения могла принимать форму позитивизма или материализма; она могла также находить выражение в форме комбинации этих двух философских направлений. Напротив, другие интеллектуалы пытались утверждать, что наука является чересчур «субъективистским» и произвольным занятием — такова, в частности, была, как принято считать, точка зрения Томаса Куна, автора нашумевшей в свое время книги «Структура научных революций», хотя быть может, сам Кун и не намеревался сказать ничего подобного. Третьи, к числу которых относятся философы-марксисты и религиозные философы, придерживаются своего рода теории двойственной истины: отдавая технические вопросы на откуп точным наукам и технике, идеологические или этические вопросы они передавали под юрисдикцию другого суда: партии, утопического будущего или церкви. Однако мало кого устраивал какой-либо из этих подходов, будь то радикальный сциентизм в его позитивистской или материалистической формах, будь то субъективизм и радикальный релятивизм, будь то какая-то разновидность теории двойственной истины. Именно это чувство неудовлетворенности и можно считать показателем того, что здесь мы сталкиваемся с реальной проблемой.

Конечно, в известном смысле эта проблема *не* является реальной. Человек, который утверждает, что этические и политические мнения не поддаются проверке, в то же самое время со страстью отстаивает свои этические и политические мнения. Юм в свое время говорил, что когда он не занимается своими исследованиями, он расстается со своим скептицизмом: точно так же поступают и релятивисты со своим релятивизмом. Однако это доказывает только то, что никто не может последовательно придерживаться релятивистской позиции. Если этим исчерпываются все те возражения, которые можно выдвинуть против релятивизма, то мы оказываемся отброшенными от релятивизма к экзистенциализму в духе 1945 года, учившему, что хотя все абсурдно, мы тем не менее должны делать выбор. А столь ли уж велика разница между ними?

Для того, чтобы зафиксировать наши идеи, давайте вспомним замечание мыслителя прошлого века, утилитаризм которого содержал в себе изрядную долю релятивизма. Вы, наверно, уже догадались, что я имею в виду Бентама и его любопытное суждение, что «если оставить в стороне предрассудки, то детская игра в «блошки» вполне равноценна искусствам и наукам музыки и поэзии». *Если оставить в стороне предрассудки, то детская игра в «блошки» сопоставима по своему значению с поэзией.*

Бентам утверждает, что предпочтение, которое мы отдаем «искусствам и наукам музыки и поэзии» перед детской игрой в «блошки», носит исключительно субъективный характер и что оно по сути дела подобно предпочтению, которое мы отдаем ванильному мороженому перед шоколадным. Бентам вовсе не собирается отрицать, что музыка и поэзия имеют большее значение, чем детская игра в «блошки» (его оговорка «если оставить в стороне предрассудки» вовсе не случайна). В контексте его утилитаризма сам тот факт, что большинство людей и *в самом деле* отдает предпочтение музыке и поэзии перед детской игрой в «блошки» делает музыку и поэзию более «полезными», а следовательно, придает им большую ценность. Однако эта ценность является, по сути дела, следствием «предрассудка», т. е. чисто субъективного интереса; относительная ценность поэзии и детской игры в «блошки» определяется не объективными обстоятельствами, но исключительно тем, что люди отдают предпочтение поэзии перед детской игрой. По сути дела, Бентам ут-

верждает, что мы отдаем предпочтение поэзии перед детской игрой не потому, что первая обладает большей ценностью, чем вторая; скорее, дело обстоит совсем наоборот, и поэзия обладает большей ценностью, чем детская игра потому, что люди отдают предпочтение поэзии перед детской игрой, *причем без видимых на то оснований.*

Обобщенный характер предложенной *формулировка* данного подхода делает его *неубедительным.* Давайте поэтому рассмотрим действительно «субъективное» предпочтение.

Одна из моделей, которую люди, по всей видимости, имеют в виду, когда говорят о субъективном предпочтении, выглядит следующим образом. Пусть существует какое-то S , которое обладает вкусом шоколадного мороженого; пусть также существует какое-то V , которое обладает вкусом ванильного мороженого. Пусть также существуют два ощущения L и D , которые означают соответственно «нравится» и «не нравится». Все, что происходит, когда говорится, что Джонсу нравится, а Смит не нравится ванильное мороженое (и соответственно, когда говорится, что Смит нравится, а Джонсу не нравится шоколадное мороженое), сводится к тому, что Джонс ощущает $V + L$, когда он ест ванильное мороженое, и $S + D$, когда он ест шоколадное мороженое. Что же касается Смита, то он ощущает $V + D$, когда ест ванильное мороженое, и $S + L$, когда ест шоколадное.

Такой подход, однако, как показал в свое время Кёлер, с психологической точки зрения является наивным. Вкус ванильного мороженого, ощущаемый Джонсом, который любит ванильное мороженое, отличается от вкуса ванильного мороженого, ощущаемого Смитом, которому не нравится ванильное мороженое. Скорее дело обстоит следующим образом: пусть то качество, которое ванильное мороженое имеет для Смита, называется V_s . V_s — это «неприятный» вкус; можно вообразить, что человек мог бы ощущать вкус V_s и он бы ему понравился, однако подобного рода случаи весьма редки и даже если бы они имели место, то были бы сопряжены со своего рода *разделением и подавлением чувств.* Короче говоря, психологически, если только не метафизически, V_s является «внутренне» неприятным. Соответственно, Смит ощущает D , когда ест ванильное мороженое, что означает, что ему не нравится его вкус, поскольку вкус ванильного мороженого обладает для него качеством V_s . Пусть

тогда V_j , то есть то качество, которое ванильное мороженое имеет для Джонса, будет внутренне «приятным», в силу чего Джонс и испытывает чувство удовольствия, когда ощущает L . Если использовать терминологию, предложенную Д. Э. Муром, то можно сказать, что вкус V_j и чувство удовольствия представляют собой единое целое для Джонса, тогда как вкус V_s и чувство неудовольствия представляют собой единое целое для Смита. С феноменологической точки зрения, их по сути дела нельзя даже разделить на две группы тем способом, который предполагается символической записью вида « $V_s + D$ », « $V_j + L$ ». Практически наверняка (за исключением специальных факторов подавления и разъединения) Смит тоже нравилось ванильное мороженое, если бы его вкус вызывал бы у него ощущение V_j , а не V_s . Точно так же Джонсу не нравилось бы ванильное мороженое, если бы его вкус вызывал у него ощущение V_s , а не V_j .

Почему в таком случае мы считаем «субъективным» предпочтение, оказываемое ванильному мороженому перед шоколадным? Иными словами, почему мы считаем его субъективным при том, что мы не относим *все* ценностные суждения к числу субъективных или соглашаемся с Бентамом в том, что если оставить в стороне наши предрассудки, то детская игра окажется столь же ценной, сколь и поэзия? Само собою разумеется, что если мы считаем все предпочтения субъективными, то мы будем считать таковым и предпочтение, которое оказывается ванильному мороженому перед шоколадным. Заслуживающий внимания вопрос заключается, однако же, в том, почему это суждение даже не кажется нам объективным до тех пор, пока мы сами не становимся Смитом или Джонсом, почему оно не обладает тем качеством объективности, которое, вне всякого сомнения, имеют многие ценностные суждения?

И дело тут не только в разногласии. Если мы полагаем, что вообще существуют какие-то объективные (или оправданные) ценностные суждения, то сходным образом мы приходим к заключению, что некоторые горячо обсуждаемые суждения являются объективно правильными. Нацисты оспаривали суждение, согласно которому *ничем не мотивированное убийство евреев из-за их расовой принадлежности является неправильным*, однако антифашисты вовсе не считают свое несогласие с нацистами по этому вопросу «субъективным». Те, кто считают, что гомосексуали-

сты должны быть полноправными членами нашего общества, страстно спорят с теми, кто считает, что гомосексуализм или гражданские права гомосексуалистов подлежат законодательному осуждению и запрещению; однако ни одна из сторон, участвующих в этом споре, не считает свою позицию чисто «субъективной». Конечно, разногласия зачастую вызывают у людей большую уверенность в своей правоте, нежели того заслуживает та моральная позиция, которой они придерживаются. Поэтому тот факт, что некоторые люди предпочитают шоколадное, а некоторые — ванильное мороженое, не оказывает решающего влияния на то, что расхождение в предпочтениях, выражаемых Смитом и Джонсом, оценивается как носящее чисто субъективный характер.

Отчасти все может упираться в то, что большинство людей не отдадут сколько-нибудь серьезных предпочтений ванильному мороженому перед шоколадным и наоборот, однако это соображение опять-таки не может играть решающей роли. Если половина населения терпеть не могла бы шоколадное мороженое, но при этом обожала ванильное, а другая половина населения терпеть не могла бы ванильное мороженое, но обожала бы шоколадное, то мы все же, — коль скоро мы бы отдавали себе отчет в своих предпочтениях, — считали бы это разногласие «делом вкуса», то есть чем-то субъективным. Поэтому существованию каких-то «нейтральных обстоятельств» нельзя отводить роль того решающего фактора, который способен разрешить интересующий нас спор.

Подлинно решающим обстоятельством является то, что самые разные биологические и психологические отличительные черты, присущие Смигу и Джонсу и обуславливающие их предпочтения, не соотносены с характерными чертами их ума и характера. Давайте устроим мысленный эксперимент и вообразим ситуацию, противоположную описанной выше, суть которой будет заключаться в том, что существует два склада характера: один, который мы считаем хорошим, как в силу его собственных достоинств, так и в силу того воздействия, который он оказывает на чувства, суждения и действия, и другой, который мы считаем плохим, как в силу его собственных особенностей, так и в силу тех последствий, которые он оказывает на поведение людей. Давайте также будем считать общеизвестным фактом,

что люди с хорошим складом характера и ума всегда оказывают предпочтение ванильному мороженому перед шоколадным, а люди с плохим характером и складом ума поступают противоположным образом. В этом случае мы, как я полагаю, поймем, что чем более мы преуспеем в осуществлении этого эксперимента, тем более мы будем склонны ощущать, что в подобного рода мире, образ которого подсказывает нам воображение, предпочтение, оказываемое ванильному мороженому перед шоколадным, будет считаться чем-то «нормальным» и «правильным», тогда как предпочтение обратного свойства будет характеризоваться как «порочное» или «недопустимое» или же как что-то подобное в том же самом роде.

Я не претендую на то, чтобы утверждать, что *любые* предпочтения *подвергаются моральной оценке* на основе тех свойств характера человека, которые они, как принято полагать, выражают. Некоторые из «предпочтений» имеют самостоятельное значение, которое весьма трудно переоценить: например, человек, который полагает, что *было бы очень забавно истязать маленьких детей ради того удовольствия, которые эти истязания доставляют мучителю, подвергся бы – если только он утверждал бы это всерьез, – осуждению уже в силу того, что он придерживается подобной точки зрения*. Однако если те вещи, которым оказывается предпочтение, сами по себе важными не считаются, *то оказываем ли мы предпочтение, исходя из содержательных оснований* или же объявляем его «делом вкуса», в общем и целом обусловлено тем, *о чем, по нашему мнению, свидетельствует это предпочтение*. *Ценностные суждения зачастую выносятся не поодиночке, но в составе какого-то множества суждений; и ряд ценностных суждений часто выражает какие-то устойчивые особенности ума, личности и характера человека*. Независимость утверждения «Я оказываю предпочтение ванильному мороженому перед шоколадным» от сколько-нибудь интересного и значимого «ряда» суждений подобного рода как раз и делает его «субъективным» (наряду, конечно же, с отсутствием какого бы то ни было внутреннего значения у этого выбора).

Даже если предпочтение, которое Смит оказывает ванильному мороженому, является «субъективным», оно от этого еще не становится иррациональным или произвольным. Смит имеет основание — наилучшее из возможных оснований — любить ва-

нильное мороженое, и кроется оно в том в том вкусе, который он испытывает, пробуя его. Ценности могут быть «субъективными» в том смысле, что они являются относительными; тем не менее это вовсе не мешает им оставаться объективными; например, то, что ванильное мороженое нравится Смиту *больше*, чем мороженое шоколадное, это — объективный факт. Айрис Мердок в своей книге «Суверенность блага» указала, что такие разные философы, как французские экзистенциалисты, с одной стороны, и логические позитивисты, с другой, на деле придерживаются одинаковой теории ценностных суждений. В основе этой теории лежит такая модель разума, согласно которой этот последний обеспечивает сознание нейтральными «фактами», отталкиваясь от которых воля должна произвольно выбирать «ценности», причем «произвольным» выбор ценностей должен быть как раз потому, что «факты» являются (по определению) нейтральными. Однако поскольку разум не дает воле никаких указаний относительно того, как она должна выносить свои суждения (согласно этой теории, разум всего лишь обеспечивает волю «фактами»), то у воли отсутствуют какие бы то ни было *мотивы* для ее произвольного выбора. Именно поэтому французские философы называют выбор «абсурдным», а философы, более склонные к натурализму, видят в инстинкте и эмоциях — исторических наследниках широко используемой категории удовольствия Бентама — главную основу морального выбора.

В том случае, который только что был рассмотрен нами выше, теория ценностных суждений, которой придерживаются экзистенциалисты и позитивисты, действует неудовлетворительно. «Факт» — то есть само *вкусовое ощущение* — и «ценность» — то есть ценные качества вкусового ощущения — это одно и то же, по крайней мере с психологической точки зрения. Данные в опыте качества как правило не являются нейтральными и зачастую требуют к себе определенного подхода или отношения. *Человек может не принимать во внимание это осязаемое требование предоставить хорошее и достаточное основание для вынесения ценностного суждения, точно так же как ребенок учится переносить боль от инъекции ради тех благ, которые приносит сделанная ему прививка, однако те prima facie благо или зло, которые несут с собой отдельные опыты, вряд ли можно отрицать.* Довольно любопытно то, что это обстоятельство нашло признание у Платона и средневековых

мыслителей; возможно, что мы — это первая культура, которая считает опыт чем-то нейтральным.

Не-нейтральность опыта имеет прямое отношение и к вопросу об относительной ценности детской игры в блошки и поэзии. Нам кажется немыслимым, что человек, который на деле ценит поэзию и понимает ее значение, который способен отличить подлинную поэзию от простого рифмоплетства и с трепетом относится к высокой поэзии, *предпочел бы* детскую игру искусствам, которые, наподобие поэзии и музыки, обогащают нашу жизнь. У нас *есть* основание для предпочтения поэзии детской игре, и это основание кроется в переживаемом нами чувстве величия поэзии, а также в тех последствиях, которые несет с собой высокая поэзия и которые находят свое выражение прежде всего в расширении воображения и чувствительности благодаря расширению набора поэтических образов и метафор и их слиянию с обыденными восприятиями и установками, которое происходит по прошествии ряда лет после того, как мы впервые познакомились с поэтическим произведением. Переживания подобного рода также являются *prima facie* хорошими или благими, и не просто благими, а облагораживающими, если использовать старомодное словечко.

Конечно же, то обстоятельство, что у ценностных суждений *могут* существовать определенного рода основания или мотивы, которые представляют для отдельных людей хорошие основания для вынесения отдельных ценностных суждений, не означает, что все ценностные суждения являются рациональными. Ценностные суждения, суждения, которые люди страстно лелеяли и во имя которых они убивали и мучили других, довольно часто основывались на отвратительной смеси агрессивных импульсов и нарциссических идей. Поэтому вовсе не удивительно, что когда такой релятивистски настроенный историк и философ, как Мишель Фуко, пишет о прошлом, он часто обращает наше внимание на подобного рода иррациональные идеи и ценностные суждения. Важно, однако же, понять, почему он поступает подобным образом.

Фуко пишет о заре нового времени (о XVI и XVII веках) и об идеологии и культуре как таковых. Его эрудиция поистине поражает, хотя многие специалисты довольно критически относятся к тем «фактам», о которых он повествует. В то время как

некоторые книги Фуко носят довольно абстрактный характер (например, «Археология знания»), другие представляют собой исследование весьма специфических проблем, например, «Рождение клиники» и «Рождение тюрьмы». «Рождение клиники» — это, пожалуй, лучшая книга Фуко, и она играет важную роль для более абстрактных теорий Фуко.

Фуко пытается показать, что «клиника», то есть больница и связанные с нею медицинские учреждения, представляют собой отражение подъема определенной идеологии относительно болезни и здоровья, равно как и следствие роста научного знания и техники. Эта идеология, в свою очередь, была связана с более широкими идеологическими изменениями, и в частности, с ростом индивидуализма в XVII веке. Со своей стороны Фуко предполагает, что, во-первых, клиника — это не лучший способ лечения больных и, во-вторых, что наша уверенность в обратном представляет собой идеологический предрассудок, иными словами, своего рода глупость или каприз.

Более далеко идущее предположение заключается в том, что идеологические убеждения и связанные с ними ценностные суждения носят произвольный характер¹. В идеологических дискуссиях невозможно занять какую-либо объективную позицию (разве что если не считать таковой таинственную точку зрения предположительно объективной «Археологии знания» Фуко).

Для того, чтобы понять, к чему ведет дело Фуко, давайте рассмотрим более привычный и менее спорный пример. В средние века было принято считать, что монархия представляет собой естественный и подлинный способ правления. Это убеждение отчасти основывалось на эмпирических убеждениях, которые в настоящее время принято считать необоснованными (одним из подобных убеждений было убеждение в том, что демократия неизбежно приводит к анархии и тирании), отчасти же — на ав-

1. Сюда же относится взгляд, что наше мышление определяется тем языком, который мы используем. Фуко говорит о «скрытых системах, которые определяют наиболее важные черты нашего поведения, причем так, что мы даже не подозреваем этого» (см.: Simon J. A Conversation with Michel Foucault // Partisan Review. № 2, 1971, p.201). Французский структурализм, по крайней мере в работах Фуко, Альтюссера, Лакана, Делеза и т. д., сводится, по-видимому, к (1) детерминизму, (2) релятивизму и (3) утверждению, что структурализм — это «лингвистическая наука».

торитете Церкви. Позиция Церкви на деле отчасти основывалась на политических соображениях (Церковь была государственной религией), однако это обстоятельство упускалось из виду, поскольку считалось, что Церковь является Богом вдохновленным и Богом установленным интерпретатором Божественного слова и Божественной воли. Со своей стороны, Фуко предполагает, что те убеждения, которых мы придерживались в недалеком прошлом, и те убеждения, которых мы придерживаемся в нынешнее время, являются не более рациональными, чем средневековая вера в божественное право королей.

Давайте на время остановимся и рассмотрим божественное право королей. Если мы не считаем, что существует *убедительное основание* для веры в существование личностного Бога, который предписывает людям, в рамках каких социальных форм и структур они должны вести свое существование, то это верование сразу же будет сочтено иррациональным, что, конечно же, само по себе не будет являться отрицанием того обстоятельства, что данное верование соответствует реальным психологическим нуждам. Даже если мы верим в Бога, в том случае, если мы не верим в то, что Церковь располагает привилегированным доступом к его желаниям, мы будем думать, что божественное право королей было иррациональным учением и что оно продолжает быть таковым и в настоящее время. Наконец, даже верующие католики будут считать, что поддержка Церковью монархии в средние века была обусловлена политическими соображениями не в меньшей степени, чем откровениями теологии. Короче говоря, вере в божественное право королей недостает и всегда недоставало адекватной рациональной основы.

Откуда в таком случае возникает соответствующее убеждение? Обычно при ответе на этот вопрос принято ссылаться отчасти на политические и экономические, а отчасти — на психологические факторы. Причем вовсе не обязательно быть марксистом для того, чтобы считать, что политические и экономические факторы оказывают воздействие на идеологию. Утешение и поддержка, обеспечиваемые верой в личного Бога и в загробную жизнь, очевидны, и столь же очевидными, вероятно, являются утешение и поддержка, которые обеспечивает вера в непогрешимость Церкви и божественно предначертанный социальный порядок. Короче говоря, нарциссическое самоудовлет-

ворение и социальные условия и были теми реальными факторами, которые вызвали к жизни это верование. Соответственно, если это убеждение является достаточно *типичным*, если оно является показательным образцом *всех* наших «идеологических» убеждений, то подобного рода факторы и в самом деле являются силами, обуславливающими *всякую* идеологию.

И именно в силу того, что значительная часть современных французских мыслителей придерживается подобного рода воззрений, они необычайно высоко оценивают идеи Маркса, Фрейда и Ницше. Маркс, Фрейд и Ницше обладают одной общей чертой: они считают лелеемые нами религиозные и этические убеждения отражением *иррациональных* факторов – в случае Маркса таким фактором является классовый интерес, в случае Ницше и Фрейда – бессознательное; Ницше к числу подобного рода иррациональных сил относит еще и волю к власти². За фасадом тех духовных и моральных представлений, которые мы склонны считать наиболее сокровенными и основополагающими, находится кипящий котел борьбы за власть, экономических интересов и эгоистических фантазий. Эта точка зрения представляет собой краеугольный камень современного релятивизма.

В то же самое время ни один релятивист сам не в состоянии использовать термин «иррациональный» так, как его только что использовал я в целях описания релятивистской позиции. Такого рода употребление данного термина исключается релятивизмом. Когда я показал этот фрагмент рукописи одному своему приятелю-релятивисту, он был возмущен моим суждением по поводу божественного права королей. Разве не был я осведомлен о том, что разумные люди под воздействием философских аргументов убедились в правильности этого учения? Предлагали ли я ничего не стоящее марксистское или же фрейдистское объяснение? Конечно же, вера в божественное право королей была *рациональной*!

На это я ему ответил, что термин «рациональный» может иметь такой смысл, согласно которому любая точка зрения, которую можно разумно и убедительно защитить, опираясь на общепринятые в какой-либо культуре положения, является «раци-

2. При этом я вовсе не собираюсь вменить в вину Марксу, Фрейду и Ницше выведение релятивистских следствий из этой позиции.

ональной», однако этот смысл данного термина не может считаться единственным или нормативно значимым смыслом. Иудеи признали Моисея законодателем и пророком, поскольку его учение удовлетворяло реальные религиозные, культурные и национальные *потребности*; однако это не то же самое, что принять учение в силу рационального доказательства. Позднее пророки миропомазали на царство иудейских царей, причем случилось этого после того, как они долго пытались убедить иудеев отказаться от института царской власти. Однако из факта миропомазания пророками иудейских царей на престол вряд ли следует, что впоследствии право христианских королей на их престол было божественного происхождения. Христианство, которое признало Ветхий Завет, стало религией Римской империи, и притом вряд ли потому, что народ или римский император обладали рациональными доказательствами его истинности. Затем римские императоры также были миропомазаны на царство, как прежде иудейские цари; однако это вряд ли может служить доказательством того, что их право на власть имеет божественное происхождение. Наконец, после принятия основных положений христианства появилась возможность приводить «рациональные аргументы» в защиту божественного права королей, *опираясь на эти основные положения христианского вероучения*. Однако пытаться выразить это обстоятельство, утверждая, что в поздней Римской империи и в средние века «вера в божественное право королей была совершенно рациональной», означает отбрасывать само понятие рациональности.

Гегель, который выдвинул идею, что Разум сам подвергается изменениям в ходе исторического процесса, использовал два понятия рациональности. Во-первых, он полагал, что рациональность поддается измерению путем сопоставления ее со стадией развития, достигнутой духом в ходе исторического процесса в данный момент времени; именно используя данный смысл понятия «рациональное», некоторые утверждают, что «вера в божественное право королей была рациональной в свое время». Во-вторых, в системе Гегеля имеется понятие рациональности как предела: оно описывает рациональность такой, какой ей предначертано стать после того, как дух пройдет все стадии своего развития. Когда современные релятивисты подвергают Гегеля «натурализации» путем отбрасывания концепции истинной

рациональности как предела, они превращают доктрину рациональности в самоопровержимый культурный релятивизм.

Ни один релятивист не желает быть релятивистом «без берегов». Как наши французские мыслители ставят границы своему релятивизму? Ответ на этот вопрос зависит от каждого конкретного философа. В случае с марксистом вроде Альтюссера предлагаемый ответ представляет собой теорию, в основе которой лежит понятие «классового интереса»: все «идеологии» являются продуктами нерациональных факторов, однако те идеологии, которые являются продуктами интересов рабочего класса, являются в нынешнюю эпоху «правильными» и ведут к освобождению человечества, тогда как те идеологии, которые возникают из интересов эксплуататорского класса, являются «неправильными» и порождают страдания. Однако Альтюссер дистанцируется от прежних сторонников классово-релятивистской точки зрения, поскольку он отказывается считать, что марксистская идеология (идеология «рабочего класса») является истинной или находится ближе к истине, чем идеология буржуазная. Согласно Альтюссеру, идеологии могут быть «правильными» и «неправильными», но они не могут быть истинными или ложными³. (Предикаты «истинный» и «ложный», утверждает он, применяются к опытным наукам, и, вероятно, к тем обыденным эмпирическим высказываниям, которые обладают ясными условиями проверки). Фуко также, по всей видимости, в большинстве своих нынешних произведений смещается к точке зрения классового интереса, хотя об этом и трудно судить с уверенностью. Суть подобной точки зрения, по крайней мере в

3. Альтюссер считает, что «философские предложения представляют собой тезисы». «Философские тезисы могут считаться — с отрицательной точки зрения — догматическими предложениями в той мере, в какой они не поддаются демонстрации в *строгом* научном смысле этого термина, то есть в том смысле, в каком термин «доказательство» используется в математике и логике, ни доказательству в *строгом* научном смысле этого термина, то есть в том смысле, в каком термин «доказательство» используется в эмпирических науках. ... Философские тезисы, поскольку они не поддаются ни демонстрации, ни научному доказательству, не могут считаться «истинными», то есть доступными для демонстрации или доказательства так, как это происходит в математике или физике. О философских тезисах можно сказать только то, что они являются «правильными»» (Althusser L. Philosophie et Philosophie Spontanee des Savants. P.: Maspero: 1967. P. 13—14.

ее радикальной альтюссерианской части заключается в том, что она пытается сбросить радикальное релятивистское положение, согласно которому ни одна «идеология» не может быть рациональной, но в то же самое время сохраняет идею, согласно которой некоторые идеологии, которым оказано предпочтение (в случае с Альтюссером это марксизм-ленинизм), могут быть хорошими, причем в основе этого утверждения лежит разделение на хорошие и плохие, «правильные» и «неправильные» идеологии, которое исходит из соображений иных, нежели соображения рациональной приемлемости. Идея состоит в том, что хотя в основе принятия всех идеологий лежат иррациональные или нерациональные причины, некоторые нерациональные причины (интересы рабочего класса) являются хорошими, и они производят (по определению?) хорошие идеологии; некоторые же нерациональные причины являются плохими и они производят плохие идеологии. Вместо того, чтобы судить идеологии по их *мотивам*, которые всегда представляют собой *рационализации*, нам следовало бы судить их по причинам.

Само собою разумеется, что этот способ ограничения релятивистской позиции не является жизнеспособным. Ведь на чем, собственно говоря, основывается суждение, что победа «интересов рабочего класса» приведет к таким явно желательным последствиям, как избавление мира от войн и расизма, а не к тоталитаризму и империализму, замаскированным под «социализм»? Если ответ на наш вопрос заключается в том, что в этом описанном нами последнем случае социализм не будет «подлинным» социализмом, что он не будет отражать «истинные» интересы рабочего класса, то на чем, с позволения спросить, основывается суждение о том, что какое-то отдельное учреждение, — например, Французская коммунистическая партия, активным членом которой является Альтюссер, — или политика будут поддерживать «истинные» интересы рабочего класса и «истинный» социализм? Если бы подобного рода убеждения можно было бы обосновать рационально, то не всякая идеология была бы иррациональной; если же эти убеждения не поддаются рациональному основанию, то любое утверждение, согласно которому какое-то учреждение или политика являются «правильными», в конечном счете должно быть столь же иррациональным, каким, как принято считать, является и любое другое «идеологическое»

утверждение. Если всякое человеческое размышление об идеологических вопросах является своекорыстной глупостью, то и идея о том, какие убеждения вытекают из «интересов рабочего класса», а какие нет, также является своекорыстной чепухой.

Возвращаясь к Фуко и оставляя в стороне те моменты в его творчестве, *которые оказываются чрезмерно радикальными*, можно сказать, что отбирая для своих размышлений те или иные события, он руководствуется *стремлением* показать нерациональность (и в действительности, иррациональность) тех реальных *мотивов*, которыми люди руководствуются, занимая ту или иную идеологическую позицию. Понятие идеологии, которое используется в данном случае, очень широко: речь идет не только о коммунизме, фашизме, демократии, божественном праве королей и т. д. Убеждения, что человек «болен» и нуждается в «исцелении», что человек «преступник» и его следовало бы, если только это возможно, «подвергнуть перевоспитанию», и множество иных наших повседневных убеждений относятся этими мыслителями к разряду «идеологических». Возможно, высказывание «если я ударю этот стакан, то он разобьется» является идеологически нейтральным, однако подобного рода высказываний встречается довольно мало.

Может сложиться впечатление, что я упустил реальную суть того, о чем говорит Фуко. Суть его размышлений, как мог бы сказать он сам, состоит не в том, что те перспективы, что задавали идеологии прошлого, были лживыми или рациональными, но скорее в том, что любая идеология в том широком смысле этого слова, в каком оно используется Фуко, не исключая и нашу нынешнюю идеологию, является культурно относительной. Он пытается показать нам, как в любой культуре образы жизни, мысли, видения, любви определяются рядом бессознательных основополагающих допущений, определяемых нерациональными факторами. Если же предыдущие идеологии кажутся нам «иррациональными», то это потому, что мы судим о них, отталкиваясь от нашего культурно обусловленного понятия рациональности.

В подходе Фуко смущает то, что определяющие факторы, на которые указывает он и другие французские мыслители, являются *иррациональными с точки зрения наших современных стандартов*. Если наша современная идеология является продуктом сил,

которые иррациональны *с точки зрения ее собственных стандартов*, то такая идеология является внутренне противоречивой. Французские мыслители не являются *обыкновенными* культурными релятивистами; они подвергают всеобъемлющей критике наше нынешнее понятие рациональности, и именно это обстоятельство бросается в глаза их читателю и смущает его.

Культурный релятивизм вовсе не является новым учением. Антропологи с самого первого момента своего существования проповедовали нам культурный релятивизм. Однако было бы ошибкой уподоблять релятивизм Фуко этому более древней форме релятивизма.

Когда антрополог проповедует нам релятивизм, то он обычно ссылается на практики и верования какого-нибудь чужестранного племени, которые с самого начала изумляют нас своей иррациональностью или омерзительностью или же и тем, и другим одновременно, и стремится показать, что эти практики и верования способствуют благосостоянию и социальной сплоченности. Короче говоря, он показывает, — в той степени, в какой приводимый им пример является разумным, — что вещи, которые в нашем обществе принято считать неправильными или иррациональными, могут быть разумными или правильными при других природных и социальных обстоятельствах.

Конечно же, ошибочное заключение зачастую выводится антропологами из их же собственных примеров, а отдельные образцы оказываются намного менее ясными, нежели привыкли считать антропологи. Очень часто антрополог будет заявлять, что все относительно, имея под этим в виду то, *что невозможно каким-либо однозначным способом определить, что является хорошим, а что — плохим*. Причем как подсказал мне Ричард Бойд, мотивы, побуждающие исследователя занять релятивистскую позицию, носят политический характер: этим шагом он пытается убедить нас прекратить разрушать архаические культуры путем критики нашей веры в превосходство рациональности и морали нашей собственной культуры. К сожалению, этот аргумент является весьма запутанным. Те примеры, которые приводит антрополог, показывают, если они являются подходящими, что определение того, что правильно, а что нет, зависит от обстоятельств, а совсем не то, что вообще не существует правильного и неправильного, пусть даже в определенных обстоятель-

ствах. Аргумент антрополога, направленный против культурного империализма, сводится к следующему: другие культуры объективно не хуже наших, поскольку согласно его точке зрения вообще не существует такого стандарта, как быть объективно хуже или лучше. Поэтому другие культуры *столь же хороши, как и наша*; поэтому неправильно уничтожать их.

Этот аргумент является двусмысленным. Заключение требует, чтобы выражение «столь же хороши, как» означало *объективно* столь же хороши (по крайней мере на наш взгляд); однако из суждения о не-существовании объективных ценностей не может следовать заключение, что одни вещи «столь же хороши» (в требуемом смысле этого слова), как и все прочие; скорее из этого суждения следует заключение, что вообще не существует такого стандарта, как «столь же хороши, как». Если бы ценности и в самом деле *носили* произвольный характер, то почему бы нам не уничтожать другие культуры, какие нам только вздумается?

К счастью, существуют более веские основания для критики культурного империализма, нежели те, что предоставляются отрицанием объективных ценностей. Мотивы, которыми руководствовался антрополог, могут быть правильными; однако он избрал ошибочную линию аргументации. Другой термин, употребляемый им двусмысленным образом, — это понятие «быть относительным». Его пример в действительности подтверждает «объективный релятивизм» Дьюи. Определенные вещи являются правильными — *объективно правильными* — при определенных обстоятельствах и неправильными — *объективно неправильными* — при других обстоятельствах, а культура и природа создают соответствующие обстоятельства. В этом антрополог прав. Однако эта точка зрения вовсе не тождественна утверждению, что ценности являются «относительными» в том смысле, что они определяются исключительно чьим-либо субъективным мнением или вкусом.

Тем не менее, *будучи освобожденным от концептуальных несообразностей*, аргумент антрополога не должен беспокоить нас. Нам следовало бы приветствовать его наблюдения, поскольку они позволяют усилить нашу восприимчивость и подвергают критике наше самоуверенное чувство культурного превосходства. Однако сама процедура сравнения аргумента Фуко с аргументом антрополога выявляет имеющееся между ними различие: Фуко не

утверждает, что прошлые практики были *более* рациональными, нежели это нам кажется; напротив, он скорее считает, что все практики являются *менее* рациональными, что в действительности они по большей части определялись неразумием и эгоистической силой. Сходство этой доктрины со старым культурным релятивизмом сильно преувеличено.

Фактом является то, что позиция, которую мы обсуждали, старается угодить интеллектуальной склонности, которая является продуктом нашего возросшего знания того, как функционируют психологические и социологические механизмы, а также нашей растущей восприимчивости к действию этих механизмов. Это знание и эта восприимчивость являются отчасти притворными, а отчасти — реальными; склонность же заключается в том, что мы попадаем в капкан, утверждая сперва, что всякий рациональный аргумент — это просто рационализация, а затем продолжая попытки *защитить* свою позицию *при помощи рациональных аргументов*.

Если бы любой «рациональный аргумент» был простой рационализацией, то не только не имело бы смысла пытаться защитить какую-либо точку зрения при помощи рациональных аргументов; было бы вообще бессмысленно *придерживаться* какой бы то ни было точки зрения. Если я считаю свое собственное согласие или несогласие с чем-либо безумным поведением, то мне следовало бы прекратить соглашаться или не соглашаться с чем-либо вообще; то, относительно чего не может быть рационального согласия или несогласия, а может быть только безумная пародия на рациональное обсуждение, вообще *нельзя* называть высказыванием. Подобно Сексту Эмпирику, который пришел к выводу о том, что его скептицизм не поддается выражению при помощи какого-либо высказывания (поскольку даже высказывание «я не знаю» не могло быть высказыванием, *известным ему*), современный релятивист, будучи последовательным, должен был бы в конце концов начать считать свои утверждения простым выражением собственных вкусов. Да и как вообще мог бы кто-нибудь быть последовательным, придерживаясь доктрины, которая делает бессмысленным само понятие «последовательность»?

Вышесказанное не означает несогласия с точкой зрения, согласно которой мы можем рационально и правильно думать о

сылки — например, абсолютизация права собственности, — *противоречат интуитивным представлениям* и не подтверждаются достаточной аргументацией. Согласно *его* точке зрения, я нахожусь в плену «патерналистской» философии, которую он считает нечувствительной к правам человека. Это — радикальное разногласие, причем оно является скорее разногласием в «политической философии», нежели простым «политическим разногласием». Однако большинство политических разногласий подразумевают разногласия в политической философии, хотя обычно они и не являются столь же сильными, как то, что было описано мною выше.

Что происходит при подобного рода разногласиях? *Если эти разногласия разумно отстаиваются обеими спорящими сторонами*, то порою все, что может произойти, сводится к тому, что одна из спорящих сторон проникательно выявляет и описывает источник разногласия. Зачастую случается и так, что коль скоро разногласия не столь фундаментальны, как то, что было описано мною выше, спорящие стороны могут модифицировать свое мнение в той или иной степени. Если на деле к согласию не приходят, то возможны компромиссные решения, которые можно классифицировать как более или менее приемлемые для той или иной стороны.

К сожалению, подобного рода благоразумные политические споры между людьми разных воззрений в настоящее время встречаются довольно редко; однако коль скоро они происходят, то доставляют нам наибольшее удовольствие. *Причем интересно то, что отношение одного человека, участвующего в споре, к своему оппоненту носит разноплановый характер*. С одной стороны, человек признает и оценивает определенные интеллектуальные достоинства наивысшей значимости: непредвзятость, готовность рассматривать соображения и аргументы, способность соглашаться с критикой, коль скоро она того заслуживает и т. д. Однако как быть с теми основополагающими положениями, по которым невозможно достичь *согласия*? Было бы непорядочно предполагать, что *в данном случае* не существует лучших и худших соображений и взглядов. Я не думаю, что позиция, согласно которой обязанность общества относиться к своим членам с состраданием, обладает приоритетом над правом собственности, определяется исключительно соображениями вкуса; с этим

согласен и мой оппонент. В том, что касается этого спорного вопроса, каждый из нас считает другого лишенным определенного рода восприимчивости и понимания. Если уж быть совершенно последовательным, то можно сказать, что мы испытываем нечто вроде *презрения*, но не к уму своего оппонента, — ведь каждый из нас проявляет наивысшую степень уважения к уму другого, — презрения не к другому как к *личности*, — ведь я испытываю больше почтения к порядочности, честности и доброте своего коллеги, чем к множеству людей, которые согласны с моими «либеральными» политическими взглядами; скорее мы испытываем обоюдное презрение к определенному комплексу эмоций и суждений, свойственных каждому из нас.

Однако вполне ли я последователен в данном случае? Я говорю, что отношусь с уважением к складу ума Боба Нозика, и это и в самом деле так. Я говорю, что отношусь с уважением к его складу характера, и это действительно так. Но если я испытываю презрение (или что-то вроде этого чувства) к свойственному ему комплексу эмоций и суждений, то разве не означает это презрения (или чего-либо подобного) к *нему самому*?

Это не так-то просто объяснить. Обычно деликатность удерживает нас от тщательного исследования того, как мы относимся к тем людям, которых мы любим и с которыми мы в то же самое время расходимся во взглядах. Фактом является то, что ни один взрослый человек не относится одобрительно и не проявляет уважения ко *всем* чертам и особенностям *какого-либо произвольно взятого человека*, не исключая и своей собственной персоны. Нет никакого противоречия в том, что какой-либо человек относится в общем и целом уважительно к другому человеку, но при этом считает, что этот последний не свободен от определенных интеллектуальных и моральных недостатков, точно так же нет никакого противоречия в том, что я в общем и целом уважительно отношусь к себе самому, но в то же самое время считаю себя не лишенным определенных интеллектуальных и моральных (или эмоциональных и т. д.) недостатков.

Я собираюсь настаивать на том, что *в мире существует разница* между оппонентом, который наделен такими основополагающими интеллектуальными добродетелями, как непредвзятость, уважение к разуму, чувство самокритики, и оппонентом, который подобного рода добродетелей лишен; между оппонентом, кото-

рый наделен впечатляющим и дельным запасом эмпирического знания, и тем, кто всего этого лишен; между человеком, который дает выход своим чувствам и фантазиям (а это свойственно всем людям, коль скоро дело доходит до того, что принято считать политическими дискуссиями) и тем, который рассуждает осмотрительно. Причем двойственное отношение уважительного презрения — это последовательное отношение: уважение мы проявляем к интеллектуальным достоинствам другого человека, тогда как презрение — к свойственным ему, — *по нашему, конечно же, мнению*, — интеллектуальным или эмоциональным недостаткам. Словосочетание «уважительное презрение» звучит почти что *непристойно*, в особенности если путать его с презрительным отношением, которое представляет собой нечто совсем иное. И отношение уважительного презрения *было бы* омерзительным, если бы «презрение», о котором идет речь, представляло собой отношение одного человека к другому человеку, а не отношение одного человека к комплексу эмоций и суждений, свойственных другому человеку. Это отношение, однако же, является намного более порядочным, нежели *ложный релятивизм*, то есть, иными словами, притязание на то, что не существует компромиссных соображений, или же различия между более убедительными или менее убедительными суждениями о предмете в ситуациях, когда человек и в самом деле чувствует, что одно мнение является разумным, а другое — иррациональным.

Может быть полезно спуститься с абстрактного уровня, на котором мы вели рассуждение, и еще раз рассмотреть относительно простой пример. Рассмотрим суждение, которое самые обыкновенные люди склонны выносить в большинстве случаев, а именно, суждение, что мир лучше, чем война. (Такого рода суждения никогда не обсуждаются Фуко, точно так же как они никогда не описываются Свифтом, и притом по одной и той же причине: оба они являются *сатириками*. Их интересует исключительно социальное безумие, а не нормальная — с социальной точки зрения — психическая жизнь, коль скоро она существует). Нет никаких сомнений относительно того, откуда ведет свое начало суждение о преимуществе мира перед войной. Мы слишком хорошо знакомы с ужасами войны, с тем, что на войне происходит со взрослыми и детьми, с военными и гражданскими

лицами, со страной и покрывающими ее растениями. Даже если своим возникновением это суждение отчасти обязано личному интересу, это вовсе *не* делает его иррациональным, совсем наоборот.

Тем не менее целые народы могут вынести противоположное суждение, согласно которому война лучше мира, причем вовсе не руководствуясь при этом соображениями о своем праве на самозащиту. Агрессивность и фантазии могут заставить целые нации жаждать чужой крови. Однако опять-таки это доказывает вовсе не то, что *все* ценностные суждения являются иррациональными, но только то, что некоторые ценностные суждения являются иррациональными; а также то, что очень трудно понять, какие ценностные суждения являются рациональными, а какие — нет в тех случаях, когда люди не в состоянии отказаться от слепой приверженности своим взглядам или подвергнуть их критике. Именно в силу этого мы придаем столь большое значение таким когнитивным ценностям, как беспристрастность и критический подход.

То обстоятельство, что некоторые ценностные суждения являются рациональными и объективными, вовсе не означает, что наш абстрактный разговор о капитализме, демократии, социализме, правах, автономии и т. д. зачастую не является бессмысленным. Даже в тех случаях, когда то, что мы намереемся сказать о подобного рода серьезных вещах, является правильным, мы зачастую имеем проблемы с адекватным выражением своих мыслей, особенно если мы не обладаем навыками и опытом выражения абстрактных идей. Случай с антропологом, который утверждает, что не существует объективных ценностей, тогда как на самом деле он имеет в виду то, что ценности обусловлены обстоятельствами, служит весьма характерной иллюстрацией для подобного рода ситуации. И даже тогда, когда мы с успехом выражаем то, что хотим сказать, *имеются* влиятельные силы нерационального толка, влияющие на наши суждения. Та цель, которую я ставлю перед собой, состоит вовсе не в том, чтобы отрицать, что сила способна испортить наши суждения, а себялюбие — совратить их. Напротив, скорее я хочу выразить свое несогласие с той точкой зрения, будто бы мы пылинки перед лицом этих влиятельных сил и будто бы пустая затея и самообман — пытаться судить рассудительно и веско. Утверждать, что

мы можем быть рациональными, вовсе не означает утверждать, что мы можем быть непогрешимыми. Напротив, как указывает Айрис Мердок, стремление к занятию логически обоснованной и рациональной позиции по сути своей представляет собой прогрессивное дело, представляет собой такое стремление, которое «способно совершенствоваться до бесконечности»⁴.

Возможно, что все сказанное мною до сих пор будет признано моральным релятивистом-интеллектуалом. Релятивисту нет необходимости отрицать рациональность всех «ценностных» суждений или же защищать точку зрения Фуко на историю, согласно которой она представляет собой дискретный ряд «дискурсов» или «идеологий», сменяющих друг друга в силу действия факторов и сил, которые не несут в себе ни грана рациональности. Более умеренный релятивист с радостью мог бы согласиться с Дьюи⁵, согласно которому некоторые ценности являются объективно относительными, то есть рационально *определенными обстоятельствами* существования, — природой и историей, — тех, кто эти ценности создает. Такой умеренный релятивист считает, что самым важным фактором является *относительность* всех ценностей. Та «объективность», которую он отвергает, — это вовсе не та объективность, которую защищал Дьюи, и *которая представляет собой просто объективность всякого суждения, подтверждаемую в его действительном экзистенциальном существовании*, но скорее та объективность, которую стал бы защищать Платон, та фиктивная, как сказал бы релятивист, объективность, которая пытается говорить с какой-то абсолютной точки зрения, *независимой* от каких бы то ни было обстоятельств и действительной для всех для них без исключения.

Если мы не согласны принять точку зрения подобного умеренного релятивизма, если мы испытываем тревогу, читая этические произведения Дьюи, то это, как я полагаю, вовсе не от того, что мы являемся мракобесами и ретроградами, поклонниками абсолютных понятий и ценностей. Когда я настаиваю на том, что убийство и пытки невинных людей морально предосу-

4. Murdoch I. The Sovereignty of Good. P. 23.

5. См.: Dewey J. The Theory of Valuation // The Encyclopedia of Unified Science. Vol. 2, № 4, University of Chicago, 1939.

дительно, я вовсе не забочусь, как я думаю, о вопросе о том, применимо ли это суждение к существу, наделенному *совершенно* иными, нежели наши, строением и психологией. Если бы, к примеру, на планете Альфа Центавра жили бы существа, которые не испытывали бы чувства боли и не имели бы никакого представления о том, что такое индивидуальная смерть, то скорее всего наше волнение по поводу «убийств и пыток» покажется им большим шумом из ничего. Однако сама чужеродность подобной жизненной формы означает, что эти существа не в состоянии понять суть тех моральных коллизий, о которых идет речь. Если наша объективность — это объективность в человеческом понимании этого слова, то она все же не перестает от этого быть объективностью.

Проблема заключается в том, что предложенная Дьюи доктрина «объективного релятивизма» не в состоянии справиться с проблемой нацизма (хотя сам Дьюи стал бы возражать против этого). Мы хотим сказать, что цели, которые нацисты ставили перед собой, были в принципе ошибочными; а утверждение, согласно которому это ценностное суждение «является истинным по отношению к нашим интересам, но ложным по отношению к интересам нацистов» представляет собой ту самую разновидность «морального релятивизма», которую мы считаем омерзительной. Объективный релятивизм кажется *правильной* доктриной для многих моральных случаев; но он вовсе не является правильным в тех случаях, когда речь идет о само собою разумеющихся правах и обязанностях и где выбор, как нам представляется, осуществляется между хорошим и плохим, добром и злом.

Конечно же, в известном смысле современное инструментальное понятие рациональности само является «объективно релятивистским». В основе этого понятия лежит дихотомия, простота которой вводит в заблуждение: суть этой дихотомии заключается в том, что выбор «целей» не является *ни рациональным, ни иррациональным* (при условии, что выполняются некоторые минимальные требования согласованности), тогда как выбор средств является *рациональным в той степени, в какой он является эффективным*. Рациональность — это характеристика средств, а не целей, и она целиком и полностью сплавлена с эффективно-

стью. Таким образом, то предпочтение, которое Джонс оказывает ванильному мороженому перед шоколадным, не является ни рациональным, ни иррациональным, однако сам акт выбора ванильного, а не шоколадного мороженого при определенных обстоятельствах был бы для Джонса рациональным, *коль скоро дана его «иерархия предпочтений»*. Эта концепция, которая восходит к утверждению Юма, согласно которому «разум является рабом страстей и должен им быть» и которая оказала громадное влияние на Бентама, нашла широкое признание в современной социальной науке. Она сыграла свою роль в экономике благосостояния и в целом ряде других областей. Современное понятие «оптимума Парето», употребляемое экономистами, представляет собой попытку создать такое понятие экономической оптимальности, которое принимает во внимание исключительно эффективность средств и не выносит никаких «ценностных суждений» о тех целях, которые преследуют субъекты экономической деятельности; это понятие вызывает настоящий интерес *как раз* потому, что оно основывается на допущении, что выбор средств поддается рациональной критике, тогда как выбор целей — нет.

Вся эта концепция, однако же, теряет значительную часть своей убедительной привлекательности, *коль скоро* мы отдаем себе отчет в том, какая чрезмерно упрощенная психологическая теория лежит в ее основе. В рамках схемы, предложенной Бентамом, цели и предпочтения рассматриваются либо как неизменные индивидуальные характеристики (иными словами, в рамках этой схемы обучение индивида изображается как процесс обучения индивида тому, чтобы лучше оценивать действительные и вероятные последствия действий и с большей эффективностью достигать поставленных целей, а не как процесс обучения индивида постановке новых целей), либо как индивидуальные характеристики, которые, не будучи в действительности *неизменными*, изменяются исключительно под воздействием факторов, которые не являются рациональными и не поддаются теоретическому объяснению.

Бернард Уильямс⁶ указал на то, что существует множество спо-

6. В данном случае я излагаю в сжатом виде лекцию «Внутренние и внешние соображения», прочитанную в Гарварде в ноябре 1978 года.

собов, при помощи которых можно подвергнуть рациональной критике не только те средства, которые индивид выбирает, чтобы достичь своих целей, но и сами цели. Эти способы становятся внятными для нас в тот самый момент, когда мы расстаемся с узким кругозором психологии Бентама и его последователей.

Психология Бентама и его последователей и в самом деле допускает *один* случай, когда индивида можно убедить отказаться от избранной им цели (или, по крайней мере, убедить отказаться от *стремления к достижению* этой цели) при помощи рациональной критики: речь идет о случае, когда индивид ошибочно оценил последствия своего поступка в том смысле, что он недооценил ту *цену*, которую ему придется заплатить за достижение поставленной перед собой цели (применительно к другим целям, у него имеющимся). Тем самым появляется повод для постановки вопроса, касающегося как воображения, так и рассудочного сознания, а именно вопроса о том, на что *могло бы быть похоже* достижение человеком какой-либо цели. Многие люди пытаются достичь целей, от достижения которых они в действительности не испытали бы особого удовольствия. Бывает и так, что то удовольствие, которое человек испытывает, оказывается не столь уж долгим или ощутимым по сравнению с тем, каким оно рисовалось ему в воображении. Даже в рамках теории Бентама и его последователей можно было бы усовершенствовать объяснение рационального принятия решений, приняв во внимание возможность неправильной оценки человеком его реальной экзистенциальной *потребности* в достижении различных целей. Вследствие этого приобретает известную степень осмысленности утверждение, согласно которому не только средства, но и цели можно подвергать критике как иррациональные.

Опять-таки, люди часто *не обращают внимания* на цели, которых они могли бы попытаться достичь, если бы задумались о них. Или же, даже если они и задумываются о них (или же кто-то указывает им на них), им может не хватать воображения для того, чтобы представить себе, к чему могло бы привести достижение этих целей, особенно в тех случаях, когда эти цели представляют собой рассчитанное на длительную перспективу формирование определенных черт и особенностей характера, например, выработку у человека способности ценить поэзию. Человек, который предпочитает детскую игру поэзии, может

быть просто не в состоянии представить себе, что означало бы для него обладание уточненным чувством восприимчивости подлинной поэзии, и если бы он мог пустить в ход свой рассудок или же предаться полету воображения, то быть может, он смог бы понять, что совершает *ошибку*.

Важно то, что способность человека рационально критиковать свои собственные цели, равно как и цели других людей, может зависеть как от его воображения, так и от его способности принимать истинные высказывания и отбрасывать ложные. Важно также еще и то, что в роли цели человека может выступать рассчитанное на длительную перспективу формирование черты ума или характера, а не какая-то вещь или событие.

Существуют еще и иные пути, — помимо ошибочной оценки человеком реального экспериментального значения своих целей или возможных альтернативных целей, — двигаясь по которым человек может совершить ошибки при выборе целей. Вильямс указывает (возрождая наблюдения, восходящие еще к Аристотелю), что очень часто цель является *общей* (например, «хорошо провести время сегодня вечером»), и проблема заключается не столько в том, чтобы найти *средства* для достижения «цели», но в том, чтобы найти общий образец деятельности, который будет конституировать приемлемое *уточнение* цели (например, «пойти на танцы» или «остаться дома и почитать книгу»). Каким будет это уточнение — творческим и оригинальным или же банальным, — это будет, опять-таки, зависеть от воображения, а не только от рассудочного сознания.

Проблема, как подчеркивает Уильямс, заключается в том, что даже если заменить психологию Бентама и его последователей с ее узким кругозором теорией, которая отдала бы дань справедливости всем вышеперечисленным вещам, нам все равно придется иметь дело с определенной формой релятивизма. В качестве примера Уильямс приводит гипотетический случай с молодым человеком, отец которого желал, чтобы его сын сделал военную карьеру. Старик апеллировал при этом к семейным традициям (на протяжении многих поколений мужчины этой семьи служили в армии в чине офицеров) и патриотизму, но напрасно. Даже когда молодой человек живо как только мог, представлял себе, что же это собственно такое — быть армейским офицером, ничто в этой цели не привлекло его. Эта цель по-

просту не была *его* целью; и вовсе не потому, что он был обделен рассудком или воображением.

Даже случай с нацистами мог бы быть в чем-то подобен этому. Предположим, что нацисты выиграли бы войну, и мы не могли бы ссылаться на поражение Германии как на практическое соображение в пользу того, чтобы не быть нацистом. Возможно, некоторые нацисты просто-напросто не представляли себе подлинные последствия нацизма, тех страданий, которые он с собой приносит и т. п. Возможно, некоторые нацисты не стали бы нацистами, если бы они были наделены рассудком и воображением, достаточными для того, чтобы оценить те последствия, которые несет с собой нацизм, или же чтобы более живо представить себе альтернативный образ жизни, образ жизни хорошего человека. Однако, вне всякого сомнения, множество нацистов все-таки осталось бы нацистами, поскольку их не волновали те страдания, к которым приводили их деяния, и поскольку как бы живо они ни представляли себе альтернативный образ жизни, он произвел бы на них не большее впечатление, нежели перспектива военной карьеры произвела на молодого человека в истории, рассказанной Бернардом Уильямсом. Они не обладают никакими целями, к которым мы могли бы апеллировать, ни действительными целями, ни целями потенциальными, то есть такими, которые они могли бы осознать, будь они наделены более способным рассудком и более сильным воображением. Даже отбросив психологию Бентама и его последователей, мы все равно сталкиваемся с проблемой морального релятивизма.

Давайте теперь рассмотрим другой, менее возбуждающий случай, нежели случай с нацистами. Представьте себе общество, состоящее из фермеров, которые по какой-то причине совершенно не интересуются искусствами, наукой (за исключением тех ее достижений, которые используются в сельском хозяйстве), религией, короче говоря, всеми духовными и культурными феноменами. (Заметьте, что я вовсе не хочу тем самым сказать, что все аграрные общества устроены или же были устроены подобного рода образом). Подобного рода людей не следует считать *плохими* людьми; если хотите, представьте их себе как совместно трудящихся, мирных, благожелательно относящихся друг к другу людей. Я хочу, чтобы ты, читатель, представил

себе, что их *интересы* ограничены такими минимальными целями, как получение достаточного количества еды и теплого крова над головой, и такими простыми удовольствиями, как совместная выпивка по вечерам. Короче говоря, представьте их себе как людей, ведущих относительно «животное» существование, и одновременно — как людей, не желающих вести какой-либо иной образ жизни.

Такого рода люди не являются *неморальными* существами. В том образе жизни, который они ведут, нет ничего непозволительного. Однако мы обычно (если только мы не исповедуем этический релятивизм) склоняемся к той точке зрения, что их образ жизни кое в чем является недостойным и предосудительным. В нем совершенно отсутствует то, что Аристотель называл «благородством». Они ведут образ жизни свиней — возможно, симпатичных свиней, но все-таки свиней, а жизнь свиньи не подходит для человека.

В то же самое время — и в этом-то и состоит вся загвоздка — мы не склонны утверждать, что люди-свиньи в чем-либо ведут себя *иррационально*. Наша позиция может быть результатом имеющего долгую историю привыкания к использованию терминов «рациональный» и «иррациональный» в духе Бентама, однако как бы то ни было, в настоящее время мы склонны именно к подобного рода словоупотреблению. Иными словами, мы хотим сказать, что жизнь людей-свиней не столь *хороша*, нежели она могла бы быть, однако они не являются *иррациональными*.

Мы не хотим сказать, что определение того, ведет ли человек лучший или худший образ жизни — это всецело дело *вкуса*. Мы не понимаем, как мы можем утверждать, что *рационально* выбирать лучшую жизнь и *иррационально* выбирать худшую. Тем не менее *отказ* от вынесения подобного рода суждений *по сути дела* выглядит как признание того, что «все относительно»; почва уходит у нас из-под ног.

Возможно, некоторые из поправок, предложенные Бернардом Уильямсом к бентамовской психологии, помогут нам справиться с *этим* затруднением. Предположим, что люди-свиньи рождаются с нормальным человеческим *потенциалом*. (Если это не так, то их образы жизни *не являются* «хуже, нежели они могли бы быть», и мы не правы, когда относимся к ним с презрением. Самое большее, что мы могли бы испытывать по отношению

к ним, — так это чувство жалости). В этом случае их можно было бы подтолкнуть к тому, чтобы они оценили художественный, научный и духовный аспекты жизни; иными словами, их можно было бы направить на более подлинный путь человеческой жизни, если можно так выразиться. И если кто-либо из этих людей пошел на это, то он, вне всякого сомнения, предпочел бы этот более подлинный образ жизни (хотя последний, быть может, не был бы столь *интересным*), нежели тот образ жизни, который он и его коллеги ведут в настоящее время. *Люди, ведущие свинский образ жизни, испытывают чувство стыда, когда начинают вести более человеческий образ жизни, тогда как люди, ведущие более человеческий образ жизни, не стыдятся того, что они делают, когда впадают в свинство.* Эти факты дают нам основание считать, что люди-свиньи допускают какую-то ошибку, своего рода *когнитивный* промах, обсуждавшийся Уильямсом. Данные сообщения дают нам основание полагать, что люди-свиньи *не обратили внимания* на альтернативные цели, а также повод для того, чтобы считать, что люди-свиньи никогда толком не представляли себе, к чему могло бы привести достижение этих альтернативных целей. Короче говоря, об этих людях нельзя сказать, что они *выбрали* худшую жизнь; ведь у них никогда не было сколько-нибудь *внятных представлений* о лучшей.

В то время как это соображение может дать нам предлог для утверждения, что образ жизни подобного рода людей-свиней поддается рациональной критике, не вполне понятно, насколько это соображение применимо к нацистам. (Можно было бы прибегнуть к обходному маневру и в тавтологической манере заявить, что любой человек, который на деле не *выбрал* лучшую жизнь, не «постиг ее адекватно»; однако подобного рода уловка вряд ли смогла бы нам помочь). Даже если брать людей-свиней, если бы они были не простыми людьми, а *идеологическими фанатиками*, то наше соображение по поводу появления у них чувства стыда могло бы не иметь силы. В этом случае просто-напросто могло бы не существовать ни одной цели, к которой мы могли бы обратиться как к такой цели, которая могла бы, по крайней мере потенциально, стать *их собственной* целью.

Наше нежелание поставить людям-свиньям в вину недостаток *соображения* (если только мы не в состоянии указать на некоторую цель, которая является, по крайней мере потенциально, *их*

собственной, и которую они не могут достичь) является результатом современных злоключений, которые, как нетрудно установить, переживает понятие разума в нашей культуре. Ни античные, ни средневековые философы не видели ничего странного в утверждении, что если *A* представляет собой *лучшую* жизнь, нежели *B*, то этот факт является *мотивом*, наилучшим возможным мотивом для того, чтобы *отдать предпочтение A*, а не *B*. Мы потеряли способность понимать то, каким образом благодать цели может сделать выбор этой цели *рациональным*.

Конечно же, всё это в значительной степени объясняется тем фактом, что мы не считаем «благо» чем-либо объективным. Однако теперь мы сталкиваемся с кругом, или скорее с двумя кривыми. Это современный круг: инструменталистская концепция рациональности поддерживает утверждение, что благодать цели еще не делает *отказ* от ее выбора иррациональным, равно как не является иррациональным выбор цели, которая является совершенно плохой. Эта точка зрения, в свою очередь, поддерживает убеждение, что благо и зло не являются объективными, а это приводит к поддержке мнения, что инструменталистская концепция рациональности является единственно разумной. Здесь мы сталкиваемся с *традиционной* проблемой: разум — это способность, которая выбирает цели на основе их *благости*, — в противоположность «страстям», которые пытаются диктовать выбор целей на основании стремлений или «склонностей». Это утверждение поддерживает мнение, что выбор блага *рационален*, а это последнее, в свою очередь, что благо и зло являются объективными. Само собою разумеется, что мы не можем взять и просто вернуться к античному или средневековому мировоззрению, как бы горячо того ни желали консерваторы; однако является ли присоединение к кругу последователей Бентама единственной остающейся нам альтернативой?

8. ВКЛАД НАУКИ В СОВРЕМЕННЫЕ КОНЦЕПЦИИ РАЦИОНАЛЬНОСТИ

Если рассмотренная нами дискуссия — а она продолжалась в течение многих десятилетий — кажется не окончательной, то это, возможно, потому, что обсуждение всегда исходит из своего рода приоритета рациональности над благом. Вопрос всегда заключается в том, можно ли в каком-то смысле назвать *иррациональным* выбор дурного — как если бы благо было под судом, а рациональность судила. Чтобы принять эту позицию, особенно когда наши предположения о рациональности представляют собой в значительной степени неисследованное собрание культурных мифов и предубеждений, мы должны заранее предрешить вопрос о статусе ценностных суждений. В оставшейся части этого исследования я предлагаю обратить термины сравнения и задаться вопросом не о том, насколько рационально благо, а о том, почему хорошо быть рациональным. Выяснение того, какую ценность имеет сама рациональность, приведет нас к прояснению природы самой рациональности и тех предположений, которые мы склонны делать о рациональности, а также позволит нам увидеть, почему неудовлетворителен тот способ, которым мы размышляли над первым вопросом.

Вспомним, что когда Макс Вебер ввел современное различие между фактом и ценностью, его аргумент против объективности ценностных суждений состоял именно в том, что невозможно установить истинность ценностного суждения так, чтобы это удовлетворило *всех возможных рациональных людей*¹. С самого начала именно невозможность, или утверждение невозможности, дать рациональное доказательство выставляла ценностные суждения в несколько подозрительном свете. Рациональность дол-

1. См. Особенно: Weber M. Die Objektivität sozialwissenschaftlicher Erkenntnis // Archiv für Sozialwiss. und Sozialpolitk. Vol. 19, 1904. P. 24—87; а также: Weber M. Der Sinn der Wertfreiheit der soziologischen und ökonomischen Wissenschaften // Logos. 7, 1917. P. 49—88, и Weber M. Wissenschaft als Beruf. Vortrag. 1919. Все три текста были перепечатаны в сборнике: Weber M. The Methodology of the Social Sciences. Glencoe, Illinois, 1949.

го судила ценность. А в этом контексте рациональность всегда означает научную рациональность²; это — результаты «положительной науки», которые считаются такими, что могут быть установлены к удовлетворению всех рациональных людей. Одна из причин для ценностного измерения рациональности очевидна. Научная рациональность, бесспорно, помогает нам в достижении различных практических целей. В то время как немного образованных людей подписались бы под тем, что мы должны развивать науку исключительно ради технологического успеха, нет сомнения, что технологический успех науки стал подавляющим в наиболее буквальном смысле слова. Мы живем в явно бесконечном ряду технологических революций — «Индустриальная

2. К. О. Апель трактует Вебера так же, как и я, в статье «Общие предпосылки герменевтики и этики: типы рациональности вне науки и технологии», *Research in Phenomenology*. IX, 1979. Апель пишет (на стр. 36):

«Макс Вебер, однако, также предложил строго отрицательный ответ на мой вопрос о возможных типах рациональности вне пределов свободной от ценностей науки и технологии. И этот ответ стал парадигматическим, как я думаю, для существующей в западном мире идеологической системы. Вебер ограничил возможности методического понимания «свободным от оценок» пониманием, которое он сосредоточил вокруг «идеального типа» «целерационального понимания» «целерациональных действий». Далее, «целерациональные действия» могут также называться «инструментальными действиями»; и в тех случаях, когда эти действия успешны, они могут быть проанализированы или реконструированы как основанные на успешных переносах правил вида «если — то» номологической науки в правила «если — то» технологических предписаний. Следовательно, Макс Вебер таким образом ограничил задачу методического понимания попыткой схватывания (свободной от ценностей) *технологической рациональности вида «средство — цель»* за пределами человеческих действий. И именно эта идея об *инструментальной рациональности* действительно составляет Веберову парадигму рациональности в ограничительном смысле.

Следует указать, тем не менее, что для *целерационального понимания* в социологии не необходимо выполнять максимальное требование удостоверения в том, что действующее лицо преуспело в переносе номологических правил на свои технологические принципы отношений «средство — цель». Чтобы понимать его действия в свете этого типа инструментальной рациональности, достаточно удостовериться, что для действующего лица было рационально действовать так, как оно действовало, исходя из понимания *своих* целей и *своих* убеждений о средствах, способах или стратегиях как подходящих для достижения его целей. Таким образом, эмпирико-герменевтической задачей понимания становится гипотетическое выяснение и проверка этих намерений цели и убеждений о средствах со стороны действующего лица, в свете которых его действия могут быть поняты как рациональные в смысле технологической рациональности вида «средство — цель».

Революция», «Электронная Революция», — которые постоянно напоминают нам о том, какую важную роль силы науки играют в формировании нашей жизни. Даже до Индустриальной Революции явно уникальный успех физики Ньютона впечатлял множество умов. Например, когда в XVII столетии впервые началось обсуждения понятия «прогресс», прогрессисты выступали с требованием, что «Ньютон знал больше Аристотеля». Никто не мог убедительно доказать, что Шекспир — лучший драматург, чем любой из древних трагиков, или лучший поэт, чем Гомер; но казалось бесспорным, что ученый Ньютон действительно и явно продвинулся в знании дальше ученого Аристотеля.

Хотя вскоре энциклопедисты и другие интеллектуалы распространили понятие прогресса с науки на политические учреждения и этику, это обобщение кажется настолько же сомнительным в XX столетии, насколько бесспорным оно казалось в XIX веке.

Огюст Конт построил свою философию — позитивизм, празднуя успех науки. История, как сообщает нам Конт, это история успеха: мы начинаем с примитивных мифов, они становятся более очищенными и утонченными, пока в конечном счете не появляются высокие религии, высокие религии в свою очередь уступают место метафизическим теориям Платона или Канта, и наконец в наши дни метафизика должна уступить «положительной науке». Очевидно, нет сомнения, кто героиня этой истории успеха: конечно — Наука. И если с самого начала меньшинство было под впечатлением от ошеломляющего интеллектуального успеха науки, нет сомнения, что на большинство произвел впечатление ее подавляющий материальный и технологический успех. Мы находимся под впечатлением от него даже тогда, когда он угрожает самой нашей жизни.

Поэтому одна из причин для сомнения в том, что ценностные суждения имеют какой бы то ни было познавательный статус, — то, что они не могут быть «проверены методами науки», как это говорилось неоднократно. Здесь также речь идет о том, что невозможно достичь всеобщего согласия, или даже согласия большинства, по этическим вопросам (как мы уже видели, этот аргумент играл главную роль в случае с Фуко). Правильен ли или неправилен аборт или гомосексуализм — это, как кажется, не те вопросы, на которые можно дать устраивающий всех ответ; тог-

да как широко распространено мнение, что правильность научной теории может быть показана ко всеобщему удовлетворению. Иногда считается, что сама рациональность науки состоит, хотя бы частично, в том, что научные предсказания могут быть предъявлены публично; что любого можно убедить в том, что это действительно заслуживающие доверия результаты, что предсказанные теорией феномены действительно имеют место. Конечно, здесь существует угроза замкнутого круга в объяснении: если мы *определяем* рациональные процедуры как ведущие к таким заключениям, по которым мы можем получить согласие большинства, то даже если правилен аргумент Вебера, состоящий в том, что в этических вопросах невозможно достичь консенсуса среди всех рациональных людей, то он будет означать лишь, что невозможно достичь консенсуса среди всех людей, использующих такие методы, которые надежно гарантируют согласие большинства или подавляющего большинства. Иными словами, мы определяем, что ценностные суждения не могут быть подтверждены к удовлетворению всех *рациональных* людей, просто замечая, что они не могут быть подтверждены к удовлетворению подавляющего большинства *всех* людей. В итоге это совсем не то, как если бы у нас был *тест* на рациональность; формулировка Вебера предполагает, что мы каким-то образом заранее пересчитали по головам всех рациональных членов общества, а затем уже выясняем, можно ли убедить их в том, что данное ценностное суждение истинно. Но ведь в действительности ничего подобного не происходит. Все примеры Вебера (о китайских мандаринах³ и т. д.) действительно показывают, что ценностные суждения не могут быть подтверждены к удовлетворению всех образованных людей или всех разумных людей, но это никак не то же самое, что всех *рациональных* людей. Ар-

3. Здесь имеются в виду неоднократно высказывавшиеся Вебером соображения о специфике «рационализации» в восточных культурах и, в частности, соответствующем типе деловой этики — конфуцианстве как идеологии китайской империи. В результате этого китайская бюрократия — мандарины — представляла собой (теоретически) меритократию, т. е. отбиралась по уровню образования. По замечанию Вебера, в то время, как в западном обществе у незнакомца спрашивали, кто его отец, в китайском интересовались, какие он сдал экзамены. См., например: Weber M. Konfuzianismus und Taoismus. Chap. 5: Der Literatenstand // Weber M. Gesammelte Aufsätze zur Religionssoziologie. Vol. 1. P. 395–430. — Прим. пер.

гумент Вебера представляет собой замаскированный аргумент мажоритаризма; он отсылает к тому факту, что мы можем добиться согласия образованных людей о «позитивной науке», тогда как мы не можем добиться их единодушия по поводу этических ценностей. Интересно сравнить эту позицию с Аристотелевой: Аристотель говорил, что, конечно, в этике мы всегда должны пытаться добиться согласия многих людей, но очень часто мы в действительности не можем этого сделать. Как ни высокомерно это прозвучит на современный слух, мы можем убедить лишь мудрого — а чтобы отличить, кто мудр, мы можем, конечно, полагаться лишь на собственное суждение.

Разумеется, в действительности мы вовсе не получаем согласие подавляющего большинства с тезисом об истинности произвольно принятой научной теории. Большинство людей прискорбно неосведомлено в науках, а многие теории, особенно в точных науках, требуют для своего понимания так много математики, что большинство людей не способны даже понять их. Хотя с этим, конечно, следует согласиться, для большинства людей это, вероятно, ничего не меняет. Согласно разбавленному операционализму, который, кажется, стал рабочей философией большинства ученых, содержание научной теории состоит в ее проверяемых последствиях, и они могут быть выражены утверждениями формы *«если мы выполним такие-то действия, то получим такие-то наблюдаемые результаты»*. Истинность утверждений такой формы, если они истинны, можно продемонстрировать, повторяя соответствующий эксперимент достаточно часто. Такой подход создает и немало проблем: проектировать, ставить и оценивать эксперименты намного тяжелее, чем может подумать обыватель. Но нет сомнения, что фактически возможно достигнуть широко распространенного соглашения по экспериментальной адекватности определенных теорий в точных науках. Принятие этих теорий обывателем может быть вопросом его уважения к экспертам, но по крайней мере эксперты, кажется, согласны друг с другом.

Конечно, в интеллектуальном отношении инструментализм не составляет сам по себе и не представляет собой надежную концепцию рациональности. Без сомнения, научные результаты имеют огромную практическую ценность; но, как мы уже сказали, ни один образованный человек не думает, что наука цен-

на исключительно своим практическим применением. И даже если бы наука была ценна исключительно своим применением, почему *рациональность* должна быть ценна исключительно своим применением? Разумеется, ценно иметь инструмент, который помогает нам выбрать эффективные средства для достижения наших различных целей; но так же ценно знать то, какие цели нам следует выбирать. Не удивительно, что истинность ценностных суждений не может быть «рационально продемонстрирована», если «рациональная проверка» по определению ограничена учреждением связей между средствами и целями. Но почему мы должны использовать в первую очередь такую узкую концепцию рациональности⁴?

Мажоритаризм также представляется непоследовательным. Разумеется, хорошо достичь соглашения о том, что считать истинным. И всегда хорошо избежать конфликта с товарищами. Но в течение столетий люди жили с неудобным знанием, что по некоторым вопросам каждый должен полагаться на собственное суждение даже тогда, когда оно отличается от суждения большинства. Многие гордились тем, что полагались на свое суждение даже тогда, когда оно отличалось от суждения большинства. Далеко не нова идея о том, что по некоторым вопро-

4. Приписывая такую узкую концепцию рациональности Веберу, Апель пишет (*op. cit.*, p. 37):

«Эта проблема Веберовой методологии «понимания» совершенно соответствовала его (более или менее неявной) философии истории. В контексте своей реконструкции истории западной цивилизации он исходил из эвристической гипотезы о том, что по крайней мере эта часть истории может быть рассмотрена как непрерывный прогресс «рационализации» и в то же время как процесс утраты иллюзий или, как он любил говорить, «разочарования» (*Entzauberung*). Под «рационализацией» он понимал усиление роли рациональности вида «средство — цель» во всех секторах социокультурной системы, особенно в сфере экономики и бюрократического управления, под постоянным влиянием прогресса науки и техники. Под процессом «разочарования», с другой стороны, Вебер понимал, среди прочего, ослабление общепринятого религиозного или философского ценностного порядка или мировоззрения. И он был готов вывести из этого развития практические следствия для своего личного мировоззрения, поскольку он утверждал, что строгий и искренний мыслитель должен принять следующую интуицию: человеческий прогресс в смысле «рационализации» дополняется отказом от идеи рациональной оценки прежних ценностей или норм в пользу обращения за помощью к окончательным пред-рациональным решениям совести перед лицом плюрализма, или, как говорил Вебер, «многобожия» прежних норм или ценностей».

сам, в частности этическим, взвешиваемые соображения настолько сложны и настолько неточны, что мы не можем надеяться на что-либо вроде научного доказательства или научного определения, но должны полагаться на восприятие и суждение. И весьма правдоподобно, что одним из высших проявлений рациональности должна быть способность судить правильно именно о тех случаях, когда нельзя надеяться «доказать» что-то к удовлетворению большинства. Действительно кажется странным, что невозможность доказать некоторые вещи ко всеобщему удовлетворению должна стать аргументом в пользу нелогичности убеждений об этих вещах⁵.

Однако, даже если эти концепции слабы в интеллектуальном отношении, дело, кажется, обстоит так, что и инструментализм, и мажоритаризм чрезвычайно привлекательны для современного сознания. Современному сознанию нравится наглядный успех; ему неудобны сами понятия суждения и мудрости. Я не социолог и не стану пытаться исследовать вопрос о том, почему индустриальное общество, и в капиталистической, и в социалистической версиях, должно восприниматься в сочетании с темами инструментального успеха и согласия большинства. Но несомненно, что социологический фактор имеет отношение к продолжающемуся увеличению роли того представления, что рациональность равняется научной рациональности, и к самой концепции научной рациональности как основанной на демонстрации инструментальных связей к (потенциальному) удовлетворению подавляющего большинства.

Концепция рациональности, которую мы только что описали, — концепция рациональности как состоящей в методах (чей характер обычно остается довольно неопределенным), которые, безотносительно к их природе, приводят к открытию эффективных связей между средствами и целями и к «публичному» утверждению этих связей, — остается в интеллектуальном отношении ненадежной, но нет недостатка в философских попытках сделать ее более представительной. Одна из этих попыток вырастает из старого эмпиризма Локка, Беркли и Юма. Ко времени Милля этот эмпиризм укрепился в то, что философы

5. В самом деле, мы видели в главе 5, что консенсусная теория рациональной приемлемости опровергает самое себя!

называют феноменализмом: эта доктрина состоит в том, что все, о чем мы в действительности можем говорить — это *ощущения*. Согласно этой концепции, даже повседневные объекты, например, столы и стулья, являются в действительности лишь множествами объективных регулярностей в действительных и возможных человеческих ощущениях.. Как выразился Милль, физические объекты — «постоянные возможности ощущения». Другой способ выразить ту же идею состоит в том, чтобы сказать, что рассуждение, которое, как кажется, идет о физическом мире, или о чем бы то ни было, в действительности выводится из описания ощущений.

Достоинство этой точки зрения, в глазах ее сторонников, состоит в том, что она позволяет им ясно говорить о том, каково содержание не только науки, но и вообще всякого познавательного значимого описания. Любая научная теория в действительности представляет собой лишь «экономный» способ заявления множества фактов формы: *если мы выполняем такие-то действия, то получим такой-то опыт*. Сторонник этого подхода не должен защищать ненадежное требование, что ученых интересует лишь применение теории или только достижение практических целей, и не интересует знание само по себе. Феноменалист не должен отрицать, что мы хотим знать природу черных дыр, что мы хотим знать, действительно ли был Большой Взрыв, каково истинное происхождение *homo sapiens*⁶ и т. д. Мы хотим знать все эти и другие вещи, потому что знание их могло бы позволять нам строить лучшие машины. Но знание этих вещей все же представляет собой лишь знание очень многих фактов формы: *если мы выполняем такие-то действия, то получим такой-то опыт* — хотя для того, чтобы показать это, требуется утонченный философский анализ. Безотносительно к нашей причине интересоваться ими, *все факты в конечном счете инструментальны*. В то же время, кажется, никак невозможно прояснить то требование, по которому назвать что-либо добродетельным значило бы сделать предсказание формы: *если мы выполняем такие-то действия, то получим такой-то опыт*. Таким образом, утверждения о добродетельности или вредности чего-либо, согласно этой концепции, не имеют никакого познавательного значения; в терминах ло-

6. Человека разумного (лат.). — Прим. пер.

гического эмпиризма двадцатого века такие утверждения вполне «эмоциональны». Однако, феноменализм сталкивается с двумя серьезными трудностями. Во-первых, требование, что утверждения о материальных объектах переводимы в утверждения о действительных и возможных ощущениях, кажется, фактически ложно. Кропотливое логическое исследование этого требования, начавшееся с работ Карнапа и Венского кружка в 1930-х годах, убедило самих феноменалистов, что это требование было необоснованным. Научные теории в целом, несомненно, заставляют нас ожидать, что у нас будет некоторый опыт, если мы выполним некоторые действия; но идея о том, что утверждения науки переводимы одно за другим в утверждения о том, какой опыт у нас будет, если мы выполним некоторые действия, была отброшена как недопустимый вид редукционизма. Во-вторых, ощущения с необходимостью носят частный характер; хотя на практике мы можем решать, действительно ли кто-либо имеет эти ощущения, просто спросив его об этом, у нас тут же возникают эпистемологические проблемы, если кто-либо задаст вопрос: «Откуда вы знаете, что этот человек связывает со своими описаниями те же ощущения, что и вы?» Если содержание науки состоит в предсказаниях о том, какие ощущения будет иметь любое рациональное существо, если выполнит определенные действия, то для того, чтобы знать, что означает само это содержание, мы должны были бы уметь определять, например, будут ли у инопланетян, если мы их встретим, те же ощущения, что и у нас, или нет и т. д. По этой причине такие философы, как Рудольф Карнап и сэр Карл Поппер, настаивали, чтобы основанные на наблюдениях предсказания науки были заявлены в форме *«если кто-либо выполнит такие-то действия, то будут иметь место такие-то публично наблюдаемые события»* – где и действия, которые будут выполнены, и наблюдаемые события, которые должны ожидаться, должны быть описаны в терминах «публичных» объектов, например, показаний измерений, а не в терминах таких частных объектов, как ощущения.

Резюмируя, можно сказать, что старый эмпиризм, или феноменализм, как казалось, обеспечивали нас точным критерием познавательной значимости: утверждение познавательно значимо, если оно переводимо в утверждение об ощущениях. Но оказалось, что любое понятие «перевода» безнадежно неопределен-

но, или же что утверждения самой науки не удовлетворяют этому критерию познавательной значимости. Трудность проведения границы между фактическими утверждениями и оценочными суждениями на том основании, что первые «переводимы» в утверждения об ощущениях, а последние — нет, состоит в том, что предполагаемая переводимость первого класса утверждений не была и, очевидно, не может быть показана. Эмпиристский редукционизм провел границу между фактическим и оценочным, но лишь ценой полного искажения картины фактического.

Наша первоначальная цель, однако, состояла в том, чтобы рассмотреть ответы на вопрос «почему добродетельно быть рациональным?» Первый ответ, который мы рассмотрели и отклонили как слишком узкий, состоял в том, что рациональность позволяет нам обнаружить надежные связи между средствами и целями. Наше внимание привлек феноменализм, потому что, если феноменализм истинен, то исчез бы очевидный конфликт между тем, чтобы интересоваться научной теорией ради ее инструментальных последствий, и тем, чтобы интересоваться теорией для изучения того, что она сообщает нам о естественных процессах. Конфликт между инструментальными интересами и вполне теоретическими интересами мог бы в некотором смысле быть удачно обойден. Конечно, все еще оставалось бы некоторое различие между этими интересами; но даже вполне теоретические интересы были бы такими фактическими интересами, которые при окончательном логическом анализе обнаружат свой инструментальный характер. *Все* знание, достойное этого названия, оказалось бы знанием связей между средствами и целями; когда связь между средствами и целями интересуется нас, потому что мы надеемся использовать ее для достижения некоторой цели, тогда мы называем наши интересы «практическими», а когда знание связи между средствами и целями интересуется нас из чистого любопытства, тогда мы называем это теоретическим интересом. Эта попытка свести все утверждения науки к утверждениям формы «если мы выполним действие А, то получим результат В», как мы отмечали, потерпела неудачу.

Заявление Карнапа и Поппера о том, что научные утверждения изложены на языке физических вещей, а не на языке ощущений, является очевидно правильным, если рассматривать его как обобщение по поводу научной практики. Однако когда оно

возводится в эпистемологический абсолют, его значение становится довольно важным. С одной стороны, если никакое утверждение наблюдения вообще не позволяет говорить о ощущениях, то интроспекция как способ научного наблюдения оказывается исключена. Хотя многие физиологи согласились бы, что она должна быть исключена, другие физиологи, не напуганные философскими и психологическими догмами, продолжают ставить эксперименты, которые опираются, по крайней мере частично, на интроспективные сообщения. Фактически, Карнап был бы не столь догматичен в запрете на интроспекцию, как некоторые физиологи-бихевиористы; он разрешил бы использование отчетов об ощущениях при условии, чтобы они рассматривались не как отчеты о наблюдениях, но как поведенческие данные, «поведение», состоящее в том, чтобы быть созданием устных сообщений непосредственно. Но что означает рассмотрение интроспективного сообщения как «вывода из вербального поведения» — в целом не ясно. Даже в случае сообщений, касающихся не ощущений, а физических объектов, например, «передо мной стол», мы обычно не принимаем сообщение, если у нас нет некоторой теории, согласно которой человек имел возможность наблюдать тот факт, о котором он сообщает. В этом смысле частью нашего общего требования последовательности нашей картины мира является то, что сами наблюдения должны быть теоретически объяснимы; если кто-либо утверждает, что наблюдал был стол в некотором месте при помощи *ясновидения*, то мы не примем это «сообщение наблюдения», потому что оно не согласуется с нашим полным корпусом теории. В этом смысле в *каждое* сообщение наблюдения входит некоторый компонент, который может быть описан как «логически выведенный». С другой стороны, когда, скажем, доктор принимает сообщение пациента о том, что тот чувствует боль, трудно знать, на основе какой «научной теории» доктор выводит из устного сообщения пациента, что тот чувствует боль; если общее предположение о том, что люди вполне могут отличить, действительно ли они чувствуют боль, может считаться теорией, то и общее предположение о том, что люди вполне могут отличить, действительно ли перед ними стол, также может считаться теорией. Трудно усмотреть какое бы то ни было фундаментальное методологическое различие между принятием чьего-то утверж-

дения, что перед ним стол, и принятием чьего-то утверждения, что он чувствует боль.

Поппер и Карнап ответили бы, что методологическое различие состоит в том, что первый вид утверждений поддается публичной проверке, а последний — нет; но оба они преувеличивают ту степень, в которой сообщения наблюдения о физических объектах всегда поддаются публичной проверке. Много таких сообщений сделано при помощи инструментов, использование которых требует длительного обучения. (Известно, что научиться «видеть» через очень мощный микроскоп можно лишь после длительного специализированного обучения и что не каждый способен к приобретению этого навыка). Принятие этой эпистемологической догмы делает ее частью определения рациональности, состоящего в том, что рациональными являются те убеждения, которые могут быть *публично* проверены. Сделать это требование частью определения рациональности очень удобно; это делает ненужным предоставление любых аргументов в пользу этого утверждения. Возможно, основание аргумента состоит в том, что вопрос о том, чт? именно не поддается публичной проверке, может вызвать разногласия, и что везде, где есть неустранимое разногласие, ничто не является правильным или неправильным. Но это означало бы принять то, что я назвал мажоритаризмом — входящую в само понятие рациональности идею о том, что все поддающееся рациональной проверке, поддается проверке, удовлетворяющей подавляющее большинство людей.

То, что логический эмпиризм по сути был сложным выражением широких культурных тенденций к инструментализму и мажоритаризму, становится очевидным, как я полагаю, в более поздней истории этого течения. Хотя логические эмпиристы отказались от феноменализма уже в 1936 году, в течение следующих двадцати лет, пока течение не начало дробиться и исчезать как распознаваемая философская тенденция, логические эмпиристы — философы науки любили говорить о «цели науки» и любили отождествлять эту цель науки с *предсказанием* (с некоторыми дополнениями и уточнениями, которые я сейчас рассмотрю). Идея о том, что цель науки — предсказание, была фундаментальной идеей позитивизма с самого его начала в работах Огюста Конта. Как мы видели, эта идея имела по крайней мере

некоторое серьезное философское объяснение, пока феноменализм был в моде; тогда можно было доказывать, что все познавательно значащие утверждения являются скрытыми предсказаниями или бесконечными множествами скрытых предсказаний. Новое появление этой доктрины, после исчезновения феноменализма, подобно появлению «примитивного» материала в ассоциациях пациента в психотерапии после того, как снята «защита». Утверждать, что целью науки является успешное предсказание (или успешное предсказание плюс нечто, описываемое как «простота»), как кажется, опасно близко к тому, чтобы утверждать, что наука преследует только практические цели; а такую позицию не стал бы поддерживать никакой философ. Действительно, философы, защищающие вполне инструменталистскую концепцию науки, делали это не потому, что являются поклонниками практики или узко мыслящими людьми, которые не могут оценить красоту абстрактного научного знания самого по себе, но скорее потому, что они чувствовали, что отождествление «познавательно существенного» с тем, что имеет ценность для создания предсказаний, поможет им раз и навсегда исключить все формы обскурантизма и метафизики. «Метафизика» была для этих философов просто другим названием для различных видов трансцендентальной спекуляции. То, чего они боялись, были религиозные и «метафизические» спекуляции (в их смысле «метафизического»).

Я считаю, что появление в культуре философской тенденции, которая была загипнотизирована успехом науки до такой степени, что не могла представить возможности знания и разума вне того, что нам нравится называть наукой — это такой феномен, которого следовало ожидать, учитывая чрезвычайно высокий престиж науки в культуре в целом и учитывая снижающийся престиж религии, абсолютистской этики и трансцендентальной метафизики. И я считаю, что высокий престиж науки *в культуре в целом* многим обязан огромному инструментальному успеху науки, вместе с тем фактом, что наука кажется свободной от бесконечных и неразрешимых споров, которые мы находим в религии, этике и метафизике.

Поскольку, однако, те профессиональные философы, которые, так сказать, рационализировали инструменталистскую тенденцию в культуре, сами не были простодушными или вполне

практически ориентированными людьми, не удивительно, что они были склонны так расширить описание «цели науки», чтобы более явно оставить место и для других целей, нежели лишь успешное предсказание. И мы находим перечисление этих других целей у логических эмпиристов в 1940-е и 1950-е годы: открытие законов, ретродикция (то есть предсказание прошлого в противоположность будущим событиям) и открытие «объяснений», под которыми эти авторы подразумевали просто дедуктивный вывод предсказаний и ретродикций из законов.

Интересно то, что здесь произошло. Чтобы подчеркнуть, что науке нужно открытие самих законов природы, а не просто предсказания, к которым ведут эти законы, эти авторы заменили простую формулу «цель науки — успешное предсказание» *перечнем* целей. Этот перечень фактически остается открытым: законы природы, оказывается, включают в себя не только законы природы в строгом смысле, то есть утверждения, которые физически невозможно фальсифицировать, но также и так называемые «законы» эволюционной теории, в действительности представляющие собой описания общих тенденций, которые могут через некоторое время, вследствие действия разумной жизни, утратить силу, и даже утверждения о носящих случайный характер диспозициях групп индивидов и даже индивидуальных организмов. Конечно, вполне правильно будет сказать, что ученые пытаются обнаружить «законы природы», включая физически случайные обобщения, которые действуют в течение длительных периодов времени и которые имеют широкое объяснительное значение, вроде тех, на которых основаны эволюционная теория или научная экономика; стремятся обнаружить существенные истины о диспозициях групп и индивидуальных организмов и организовать все это в дедуктивные (и индуктивные) структуры. Но почему этот перечень именно таков?

Перечень задуман таким образом, чтобы охватить все виды истин, которые ученые стремятся обнаружить (конечно, в физической науке) и при этом не включать никакого нежелательного («познавательного незначимого») материала. Прежний поиск «критерия познавательной важности», вроде «предложение значимо, если и только если его можно подтвердить или фальсифицировать», был заменен перечнем типов утверждений — таким, что утверждение должно быть принято, если оно при-

надлежит к одному из этих типов, а иначе должно быть отклонено. Но почему вообще для философа приемлем такой ход? Даже если истинно, что все утверждения дисциплин, которые мы называем «науками», принадлежат к этим типам (а вовсе не ясно, что дело обстоит именно так — является ли историческое объяснение действительно лишь подведением ретродикций под «законы?»), следует ли из этого, что верификация этих типов утверждений, и только этих типов утверждений — цель самого разума, а не цель лишь специальных применений разума, которые мы называем науками? Ответ, конечно, состоит в том, что эти философы нисколько не сомневались, что «наука» исчерпывает разум. Но почему они не сомневались в этом? Они не сомневались относительно этого, потому что для них речь шла не о противопоставлении науки, в смысле знания, оперирующего по существу методами эмпирических и математических наук, неформальному разуму, оперирующему методами, которые могли бы быть приспособлены к интересам, отличным от интересов наук, но не в меньшей степени могут отвечать настоящим стандартам. Скорее они имели в виду противопоставление знания, оперирующего научными методами, и псевдознания, притворяющегося оперирующим откровением или какими-либо забавными нисобыкновенными способностями. Разум должен быть коэкстенсивен с наукой, потому что *А Что Же Еще?* Однако утверждение этих философов, что разум и наука — это одно и то же, поставило их в затруднительное положение. Например, поскольку они не хотели отрицать, что есть такая вещь, как историческое знание, они отстаивали точку зрения, согласно которой история является наукой, и даже точку зрения, согласно которой работа историка в действительности состоит в подведении индивидуальных утверждений о прошлом под законы — а это суждение об истории кажется вполне ложным.

Возможно, неудивительно, что течение логического эмпиризма начало распадаться к 1950 году. Мы взглянули на это течение исключительно с точки зрения одного вопроса, хотя у логических эмпиристов было очень много различных философских интересов, и они получили много ценных результатов. Однако, с точки зрения нашего вопроса «насколько хороша рациональность?», течение логического эмпиризма представило аргументированную философскую защиту той точки зрения, что ответ,

причем единственный, на наш вопрос состоит в том, что рациональность хороша для открытия связей между средствами и целями. Философская доктрина феноменализма обеспечила логических эмпиристов интересной философской защитой этого требования. Когда феноменализм был отброшен, а философская *защита* требования сменилась самим требованием, и даже более того, когда само требование было сделано «более разумным», допускающим исключения, модификации и т. д., то вся могучая сила этого течения исчезла. Позиция, заключающаяся в том, что цель самого разума — открытие предсказаний, ретродикций, законов природы и их систематизация, и это — *все* цели разума, создает ту проблему, что нет просто никакой причины верить этому (я не хочу сказать, что есть причина полагать, что это ложно). Если понятие закона природы расширено таким образом, чтобы открытие законов природы включало открытие диспозиционных утверждений об индивидуальных организмах, а понятие диспозиции настолько широко (или настолько неопределенно), что утверждение, что некоторый ученый завидует репутации своего коллеги, считается «диспозиционным» утверждением, а утверждение, что этот ученый рассказал некоторый анекдот потому, что он ревновал к репутации своего коллеги — «подведением конкретного события под закон», то, может быть, все, что кто-либо говорит, может интерпретироваться или как заявление общих законов, или как подведение описаний под общие законы. Возможно, даже высказывание о ком-либо, что он является *нравственно хорошим*, может быть рассмотрено как приписывание этому человеку «диспозиции». Нет, проблема с попыткой определить цели познавательного исследования вообще посредством перечня состоит в том, что сам перечень должен быть *истолкован*: если термины перечня истолкованы более или менее буквальным способом, то в утверждения перечня даже не входили бы все виды утверждений, в обнаружении которых заинтересованы ученые, если под «ученым» мы подразумеваем *и историка, и психиатра, и социолога*; в то время как, если термины перечня истолкованы настолько мягко, что нет никакой трудности в рассмотрении утверждений, сделанных историками (и описательных утверждений на языке повседневной психологии), как принадлежащих к типам, включенным в перечень, то перечень становится ничего не стоящим.

В любом случае, в отсутствие любого эпистемологического объяснения того, почему утверждения *этих видов* и только утверждения этих видов должны быть способны к рациональной проверке, такой перечень был бы не более чем гипотезой о пределах рационального исследования. Простая гипотеза, будь то в форме перечня или в некоторой другой форме, не могла бы иметь исключающей силы, которой логические эмпиристы требовали от «критерия познавательной значимости».

ФЕТИШИЗМ «МЕТОДА»

Поскольку ответ на вопрос «почему хорошо быть рациональным?» не может состоять просто в том, что рациональность позволяет достигать практических целей, и не может состоять просто в том, что рациональность позволяет обнаружить связи между средствами и целями, мы можем рассмотреть другой возможный ответ, который в разное время выглядел довольно привлекательно. Многие философы считали, что наука следует определенному *методу*; если фактически у нас есть метод с таким свойством, что, используя этот метод, можно надежно обнаруживать истины, а никакой другой метод не имеет никакого реального шанса на обнаружение истин; и если экстраординарный успех науки (и постоянные противоречия в других областях, нежели наука) *объясняется* тем, что наука, и только лишь наука последовательно использует этот *метод*, то, возможно, рациональность (в той степени, в какой она вообще существует) *должна* быть идентифицирована с владением этим методом и его использованием. Ответ на вопрос «почему хорошо быть рациональным?» состоял бы тогда в том, что рациональным хорошо быть потому, что рациональный человек может обнаруживать истины (любого вида, какие ему нужны), тогда как нерациональный человек не имеет никаких реальных шансов на обнаружение истин, кроме слепого везения. Подобно инструментальному подходу, этот подход имел свою философскую историю взлета, застоя и упадка. От публикации «Логики» Милля в 1840-х до публикации «Логических оснований вероятности» Карнапа, влиятельные философы науки продолжали считать, что в основе эмпирической науки лежит нечто вроде формального метода («*индуктивная логика*»), и что длительная работа могла бы

привести к явному утверждению этого метода, к формализации индуктивной логики, сопоставимой с формализацией дедуктивной логики, которая была достигнута, начиная с появления в 1879 году работы Фреге. Если такой метод *был бы* обнаружен, то даже если это само по себе не *доказывало бы*, что рациональность исчерпывается этим методом, тем не менее бремя доказательства было бы возложено на людей, считающих, что существуют истины, рациональная приемлемость которых может быть обоснована или показана любым другим методом.

Согласно наиболее влиятельной школе, так называемой «Бейесианской» школе, общий характер этого индуктивного метода, который пытались формализовать философы, таков: мы принимаем или притворяемся, что язык науки формализован и что ученые имеют доступ к некоторому числу надежных наблюдений, выражимых «предложениями наблюдения» на этом формализованном языке. Мы также предполагаем, что различные рассматриваемые гипотезы выражены формулами этого языка. Проблема индуктивной логики принимается за проблему определения «функции подтверждения», которая является функцией вероятности, определяющей математическую вероятность каждой из гипотез, основанных на данных наблюдения, или, иначе говоря, «степень поддержки» подтверждениями каждой из альтернативных гипотез. Обычно мы предполагаем, что знаем вероятность того, что данное подтверждение было бы получено, если бы каждая из альтернативных гипотез была истинна; это так называемая «априорная вероятность», то есть вероятность получения подтверждения для *данной* гипотезы. Что мы хотим вычислить — это так называемая «апостериорная вероятность», которая является вероятностью гипотезы при *данном* подтверждении. Теорема Бейеса дает эту «апостериорную вероятность» как простую функцию априорных вероятностей⁷ и не

7. Классическая теорема Бейеса такова: пусть $A_1, A_2, \dots, A_n$ — некоторые попарно несовместимые события, хотя бы одно из которых обязательно наступает, и B — некоторое событие. Тогда условная вероятность A_k при условии, что наступило B , может быть определена по формуле:

$$P(A_k | B) = \frac{P(B | A_k)P(A_k)}{\sum_{k=1}^n P(B | A_k)P(A_k)}$$

где $P(A_k)$ — априорная вероятность, $P(A_k | B)$ — апостериорная. — Прим. пер.

которых других вероятностей, так называемых «предшествующих вероятностей» альтернативных гипотез, то есть вероятностей или «субъективных степеней убеждения», назначенных учеными этим альтернативным гипотезам до исследования наблюдательного подтверждения.

«Априорные вероятности» действительно легко вычислить в двух наиболее общих случаях: их легко вычислить, когда (а) гипотеза фактически подразумевает подтверждение (в этом случае «априорная вероятность» подтверждения для данной гипотезы равна *единице*); или когда (b) сама гипотеза является статистической или стохастической гипотезой, частью содержания которой является то, что конкретное полученное подтверждение должно произойти с некоторой вероятностью p . Трудность в применении теоремы Бейеса — трудность настолько серьезная, что и философы, и статистики глубоко расходятся по вопросу о важности и полезности теоремы Бейеса для подтверждения теорий — состоит в необходимости предварительной системы измерения вероятности, множества «субъективных степеней убеждения», в терминологии де Финетти и Сэвиджа⁸.

Ограничимся сейчас такими гипотезами, для которых действительно можно вычислить «априорные вероятности». Для гипотез этого вида описанный выше метод является действительно вполне формальным методом; то есть мы могли бы запрограммировать вычислительную машину так, чтобы вычислять степени поддержки различных гипотез при соответствующих «входных» данных. Но эти входные данные должны были бы включать не только вычисляемые «априорные вероятности», но также и измерения предшествующей вероятности в данном контексте. Если мы думаем об этих измерениях предшествующей вероятности как о представлении предшествующих убеждений ученых о мире, как предлагает термин «субъективная функция вероятности», то получается, что во входные данные для самого метода включено множество независимых фактических убеждений (или степеней убеждения) о мире. Таким способом сегодня рассматривают этот вопрос многие философы науки; становится все более и более распространенным мнение, что невоз-

8. См., например: *Savage L. J. The foundations of statistics*. N. Y.; L., 1954. — Прим. пер.

можно провести четкую границу между *содержанием* науки и *методом* науки; что метод науки фактически постоянно изменяется по мере того, как изменяется содержание науки. Теорема Бейеса, если она действительно схватывает логику подтверждения теории, обеспечивает способ формализации этой зависимости метода науки от содержания науки через потребность в функции предшествующей вероятности.

Чтобы поставить этот вопрос несколько более абстрактно, мы могли бы сказать, что фетишист «метода» считает рациональность *неразделимой*. Но теорема Бейеса указывает, что дело обстоит не так, что мы можем разделить рациональность даже в области науки, даже в специальной области, где мы имеем дело с теориями, для которых априорные вероятности вычислимы, на две части: *формальную* часть, которая может быть схематизирована математически и запрограммирована на компьютере, и *неформальную* часть, которая не может быть так схематизирована и которая зависит от действительных меняющихся убеждений ученых. И было бы, мягко говоря, чудесно, если бы формальной части рациональности было достаточно, чтобы гарантировать хорошие результаты. Тогда мы могли бы сказать, что если ученые проводят наблюдения тщательно, собирают достаточно наблюдений и вычисляют степени поддержки по теореме Бейеса, то в конечном счете они придут к соглашению, даже если бы они не соглашались вначале, вследствие различия в их субъективных степенях убеждения. Но в этой счастливой картине есть две ошибки.

Первая состоит в том, что даже если бы мы могли показать, что в конечном счете «предшествующая функция вероятности» сходится на нет или что ученые с различными предшествующими функциями вероятности в конечном счете приходят к соглашению, если продолжают собирать большее количество подтверждений и использовать теорему Бейеса, все же еще потребовалось бы, чтобы эта сходимость была разумно быстрой. Если ученые с различными предшествующими функциями вероятности не придут к соглашению, пока не произойдет то явление, которое должно быть предсказано, или пока не пройдут миллионы лет, то на любом промежуточном этапе наличие некоторой математической гарантии возможной сходимости бесполезно; проблема с обоснованиями «в конечном счете» состоит в том,

что «конечный счет» может быть слишком отдален. По известному выражению Джона Мейнарда Кейнса, «в конечном счете все мы будем мертвы». Вторая ошибка состоит в том, что, как оказывается фактически, различия в функции предшествующей вероятности могут вести к сильным различиям в действительных степенях поддержки предоставляемой теории и что эти различия могут составлять то, что обычно рассматривалось бы как явная иррациональность.

Последнее замечание можно выразить другим способом: ученый назначит степени поддержки гипотезам, выглядящим «разумными», только если он отправляется от «разумной» функции предшествующей вероятности. Если человек только повинуется формальной части описания рациональности, если он логически последователен и назначает степени поддержки в соответствии с теоремой Бейеса, но его функция предшествующей вероятности крайне «неразумна», то его суждения о том, до какой степени различные гипотезы поддержаны подтверждением, будут (как ученые и обычные люди в действительности судят об этих вопросах) чрезвычайно «иррациональны». Формальная рациональность, обязательство к формальной части научного метода, не гарантирует реальную и действительную рациональность.

Поразительно, до какой степени это верно. Артур Бёркс фактически показал, что есть даже «контр-индуктивные функции предшествующей вероятности»⁹. Это означает, что логически возможно такое измерение предшествующей вероятности, что, если бы у ученого было это измерение, то чем больше подтверждений получала бы гипотеза (на основе наших обычных индуктивных суждений), тем меньшую вероятность ученый назначал бы этой гипотезе в течение очень долгого времени.

Одним из способов избежать трудности может быть попытка дополнить существующую формальную теорию научного метода дальнейшим набором формальных правил, которые определяют, какие из функций предшествующей вероятности являются разумными, а какие — неразумными. Но нет серьезных оснований полагать, что возможен набор правил, которые могли бы

9. См., например: *Burks A. W. Chance, Cause and Reason: an Inquiry into the Nature of Scientific Evidence. Chicago, 1977. — Прим. пер.*

отличить разумные функции предшествующей вероятности от неразумных и которые будут проще, чем полное описание всей совокупной психологии идеально рационального человеческого существа. Надежда на формальный метод, способный к абстрагированию от фактических человеческих суждений о содержании науки (то есть о природе мира) и от человеческих ценностей, кажется, испаряется. И даже если мы расширяем понятие метода так, чтобы формализация психологии идеально рационального человека-ученого считалась бы «методом», нет никакой причины думать, что «метод» в *этом* смысле был бы независим от суждений об эстетике, суждений об этике, суждений о чем угодно. В конце концов, причина считать, что научный метод не обращается к убеждениям об этическом, эстетическом и т. д., и не предполагает их, как раз и состоит в убеждении, что научный метод — это *формальный* метод.

Мое обсуждение зависело от принятия правильности одного специфического подхода к формализации научного метода, так называемого бейесианского подхода. Но подобные проблемы возникают и с каждым из других предпринятых подходов. Даже если мы попробуем изолировать некоторую маленькую часть индуктивного метода, которая не была бы столь же грандиозна, как подтверждение теорий, но которая находилась бы в большем соответствии с тем, что Бэкон понимал под «индукцией», то есть даже если мы попробуем выделить метод для подтверждения простых обобщений, исследуя достаточное число случаев и «проецируя», или «отображая», истинность обобщения, то возникают подобные же проблемы. Нельсон Гудмен¹⁰ показал, что ни одно *вполне формальное* правило индуктивного отображения не может быть свободно от противоречий; прежде, чем формальное правило может даже предположительно выдать последовательные результаты, надо *заранее* разделить предикаты языка на те, которые мы хотим расценивать как «отображаемые», и те, которые мы будем трактовать как «неотображаемые»¹¹. Даже наиболее элементарная часть индукции, оказыва-

10. См.: *Goodman N. Fact, Fiction and Forecast*. 2nd ed. Hackett, 1977. Первое издание появилось в 1954 году. [Ср. русский перевод в кн.: *Гудмен Н. Факт, фантазия и предсказание* / Пер. с англ. А. Л. Никифорова // Гудмен Н. Способы создания миров. М.: Идея-Пресс, Праксис, 2001. — *Прим. пер.*].

11. В русском переводе А. Л. Никифорова (см. предыдущее примечание) для

ется, включает неформальную часть (а именно разделение словаря на отображаемую и неотображаемую части), и это вновь настоятельно поддерживает предложенное в нашем обсуждении теоремы Бейеса заключение о том, что нельзя провести четкую границу между фактическими убеждениями ученых и научным методом. Для этого Гудмен изобрел предикат «зелубый», который применим к вещам в том случае, если они *наблюдались до 2000 года и зеленые или не наблюдались до 2000 года и голубые*. До 2000 года все, что подверглось исследованию и было определено как зеленое, также является исследованным и зелубым. Любое формальное правило отображения, которое сообщало бы нам, что когда мы исследовали некоторое число вещей, скажем, изумрудов, со свойством P , то мы можем вывести, что «все изумруды — P », разрешило бы нам делать противоречащие выводы, что «все изумруды являются зелеными» и «все изумруды — зелубые». И Гудмен убедительно показывает, что все попытки исключить «причудливые» предикаты, подобные «зелубому» на вполне формальных основаниях не будут работать¹².

Есть действительно близкая связь между трудностью Гудмена

передачи термина «projectability» не вполне удачно употребляется термин «экстраполяция». Экстраполяция — это приближенное определение значения функции для точек, лежащих вне определенного интервала, по значениям в точках, находящихся внутри этого интервала; Гудмен говорит не об этом. Отображение (как и *projection* по-английски) — термин теории множеств, означающий наличие определенного соответствия между элементами одного множества и другого. Понятие отображения является ключевым при обсуждении проблематики изоморфизма и теоремы Цермело, в том числе и для Гудмена, предложившего критерий «экстенционального изоморфизма» (для чего им и вводится понятие «отображаемости») в «Структуре явления» (*Goodman N. The Structure of Appearance. Cambridge Mass., 1951*) — более крупной и специальной работе, к которой я отсылаю любознательного читателя. См. также мою статью: *Лебедев М. В. Миротехнологии Гудмена // Гудмен Н. Способы создания миров. М.: Идея-Пресс, Праксис, 2001. С. 363–376. — Прим. пер.*

12. Собственное решение Гудмена состоит в том, чтобы рассмотреть форму и историю предшествующего отображения предикатов, участвующих в выводе (наряду с некоторыми связанными вопросами, например «закрепленность» (*entrenchment*) и «отвергание» (*over-riding*). По предложению Гудмена отсюда следовало бы, что культура, которая *всегда* отображала такие «сумасшедшие» предикаты, как его знаменитый предикат «зелубый», была бы теперь совершенно обоснована в этом — ее выводы были бы теперь индуктивно обоснованы!

В то время как я соглашаюсь с Гудменом, что соответствие с прошлой практикой — важный принцип в науке, Гудменова версия этого принципа слишком проста и слишком релятивистична.

с Бэконовой индукцией и потребностью в функции предшествующей вероятности в связи с теоремой Бейеса. Предположим две гипотезы, между которыми ученый должен выбрать (в некоторое время до 2000 года) — «все изумруды — зеленые» и «все изумруды — зелубые». Предположим, что релевантное подтверждение состоит в том, что очень многие изумруды были исследованы и сочтены зелеными (и, следовательно, все также и зелубыми). Если ученый вычисляет степень поддержки для этих двух гипотез, используя теорему Бейеса, то оказывается, что он может обнаружить или намного более высокую степень поддержки для нормальной гипотезы («все изумруды — зеленые»), или намного более высокую степень поддержки для ненормальной гипотезы («все изумруды — зелубые»), или равную степень поддержки для обеих гипотез, в зависимости от своей функции предшествующей вероятности. Если измерение чьей-либо субъективной вероятности назначает намного более высокую предшествующую вероятность утверждению «все изумруды — зеленые», чем утверждению «все изумруды — зелубые», тогда этот субъект будет, фактически, вести себя, как если бы он отображал предикат «зеленый» и не отображал «зелубый». С бейесианской точки зрения потребность в решении о том, какие предикаты являются отображаемыми, а какие — нет, до того, как начать индукцию — лишь частный случай потребности в функции предшествующей вероятности.

Карл Поппер предложил, что нужно принимать *наиболее фальсифицируемую* из альтернативных гипотез; но оказывается, что его формальные меры *фальсифицируемости* дадут различные результаты, в зависимости от того, какие предикаты языка мы принимаем за примитивные. Считаем ли мы ученого, вслед за Поппером, пытающимся найти *наиболее фальсифицируемую гипотезу, которая еще не была исключена*, или считаем, более традиционно, что мы пытаемся вычислять *степени поддержки для гипотез*, у нас все еще остается необходимость в неформальном элементе, соответствующем Гудменову решению о том, что некоторые предикаты отображаемы, а другие — нет, или соответствующем принятию Бейесовой функции предшествующей вероятности.

Здесь читатель может задаться вопросом: если никакого научного метода не существует или если метод, насколько он может быть формализован, зависит от неформализуемых исход-

ных данных, тогда как мы объясняем успех науки? Бесспорно, что наука была и остается поразительно успешным предприятием. Мы склонны чувствовать, что причина для ее успеха должна иметь некоторое отношение к различиям между способами, которыми ученые приобретают знание, и способом, которым люди традиционно приобретали знание в донаучные эпохи. Полностью ли это неправильно? Ответ будет отрицательным. Альтернативы, между которыми мы должны выбирать, состоят не в том, что наука преуспевает, потому что следует некоторому строгому формальному алгоритму, или наоборот, преуспевает по чистой случайности. Начиная с XV столетия и с кульминацией в XVII веке, ученые и философы начали выдвигать новый набор методологических принципов. Эти принципы — не строгие формальные правила; они требуют применения неформальной рациональности, то есть интеллекта и здравого смысла; но, однако, они формировали и продолжают формировать научное исследование. Короче говоря, научный метод существует; но это предполагает предшествующие понятия рациональности¹³, это — не метод *de novo*, который мог бы служить всеобъемлющим и окончательным методом, давая самое определение рациональности.

Одним из ведущих методологов XVII столетия был физик Бойль. До XVII столетия физики не проводили резкого различия между фактическим выполнением экспериментов и простым описанием мысленных экспериментов в подтверждении теорий, которые они принимали на более или менее априорных основаниях. Более того, физики не видели необходимости в обнародовании описаний неудачных экспериментов. Короче говоря, эксперименты понимались в значительной степени как иллюстрации к доктринам, принятым на дедуктивных и априорных основаниях, а не как свидетельства за и против теорий. Бойль написал руководства по процедуре экспериментирования, подчеркнул необходимость проводить четкое различие между мысленными экспериментами и фактическими экспериментами, а также необходимость исчерпывающего описания

13. Сам Милль соглашается с этим (примечательно ворчливым тоном), когда пишет, что мы не можем ожидать, что индуктивный метод будет работать, «если мы предполагаем, что он будет сочетаться со всеобщим идиотизмом» («Утилитарианизм», глава 2).

всех выполненных экспериментов, особенно тех, которые потерпели неудачу. Сам Бойль был учеником философа Фрэнсиса Бэкона, и несомненно, пришел к оценке важности этих правил через индуктивную перспективу Бэкона; фактически, однако, конкретные инструкции, данные Бойлем, возможно, были более важны или столь же важны в формировании направления физических исследований, как и более абстрактная и схематическая защита индуктивной процедуры, данная Бэконом.

Отказ от попыток устанавливать теории априорно в пользу попыток проверять теории, выводя из них проверяемые заключения и выполняя эксперименты, конечно, был методологическим сдвигом. Как мы видели, однако, мы не можем просто отождествить «*быть рациональным*» с «*принимать теории исключительно на том основании, что они поддержаны тщательно выполненными экспериментами*». Так, даже в науке не всегда возможно провести управляемые эксперименты. Иногда приходится полагаться на пассивное наблюдение скорее, чем на какой-либо вид активного вмешательства, которое подразумевается термином «эксперимент». И, как мы уже видели, даже когда мы тщательно выполним эксперименты с целью выбора среди альтернативных теорий, оценка той степени, до которой экспериментальные результаты поддерживают различные альтернативные теории, по-прежнему остается полностью неформальным вопросом.

В противовес тому подходу, который мы сейчас выдвинули, Карл Поппер неоднократно доказывал, что определенный научный метод *существует*, что он может быть описан и что мы должны полагаться в раскрытии природы мира только на него.

Поппер думает, однако, что есть и более широкие понятия рациональности, чем научная рациональность, которые применимы и к принятию этических решений.

В своей концепции, сформулированной в его влиятельной книге «Логика научного исследования» и в последующих публикациях, Поппер доказывал, что научный метод состоит в выдвижении «в высшей степени фальсифицируемых» теорий; теорий, которые дают рискованные предсказания. Тогда мы продолжаем проверять все теории до тех пор, пока не останется только одна. Тогда мы принимаем выжившую гипотезу как допустимую в настоящее время и повторяем всю процедуру. Поскольку устранение всех теорий, кроме одной, проводится на

дедуктивных основаниях — теория устранена, когда она подразумевает такое предсказание, которое определенно фальсифицировано — использование теоремы Бейеса не требуется, и никакая оценка степеней поддержки не имеет к этому отношения, как утверждает Поппер.

Одна из проблем с подходом Поппера состоит в том, что невозможно проверить все настоятельно фальсифицируемые теории. Например, теория, состоящая в том, что если я надену на голову мешок из-под муки и постучу по столу 99 раз, то появится демон, будет сильно фальсифицируема, но я, конечно, не собираюсь потрудиться проверить это. Даже если бы я захотел проверить это, я мог бы подумать о 10^{100} подобных теориях, и продолжительности человеческой жизни, или даже продолжительности жизни человеческого рода, не хватит, чтобы проверить их все. В таком случае, по логическим причинам необходимо выбрать, на методологических основаниях, очень небольшое количество теорий, которые мы действительно позаботимся проверять; и это означает, что нечто вроде предшествующего выбора присутствует даже в попперианском методе. Как я отметил выше, даже вычисления Поппером степеней фальсифицируемости находятся в зависимости от того, какие предикаты мы рассматриваем как примитивные в нашем языке, и в этом смысле даже понятие фальсифицируемости требует предшествующего решения, аналогичного решению Гудмена о том, что некоторые предикаты «отображаемы», а другие — нет. Однако оставим эти технические моменты, которые в любом случае не представляют для нас интереса в настоящем обсуждении. Даже если попперианский метод неполон и должен быть дополнен более интуитивным методом, который мы в настоящее время не способны формализовать, не может ли быть так, что он описывает необходимое — если не достаточное — условие научной рациональности? Разве не может быть, короче говоря, что необходимым условием для приемлемости научной теории является то, что она прошла попперианскую проверку? Сама же попперианская проверка может привлекать предшествующий набор теорий, чтобы проверить, что? именно является неформальным и для чего у нас нет алгоритма. Чтобы вычислить, какие теории сильнее всего фальсифицируемы, нам могут понадобиться неформальные решения, для которых у нас

нет алгоритма; но мы все же можем настаивать, чтобы никакая теория не была принята, если сначала не отобрано множество *интуитивно* «в высшей степени фальсифицируемых» теорий и если все эти теории — кроме той, которую мы принимаем, — не опровергнуты впоследствии тщательно проведенными экспериментами. Короче говоря, разве мы не можем посоветовать ученому: действуйте, как советует Поппер, а там, где методы Поппера невозможно формализовать, полагайтесь на свою интуицию о том, как должны интерпретироваться попперианские принципы? И разве не может быть так, что попперианский метод, неопределенный и неформальный (каков он частично), исчерпывает не только понятие *научной* рациональности, но и всей *познавательной рациональности*, то есть разве не может быть так, что утверждение является обоснованно утверждаемым или рационально приемлемым, если и только если оно следует из теории, которая может быть принята на основе попперианского теста? Ответ состоит в том, что такая концепция рациональности является слишком узкой *даже для науки*. С одной стороны, это исключило бы принятие одной из наиболее успешных и широко приветствуемых из всех научных теорий, Дарвиновой теории эволюции путем естественного отбора. Это — следствие, которое сам Поппер старается принимать хладнокровно, но, конечно, научное сообщество не может мириться с этим. Теория естественного отбора не является высоко фальсифицируемой; она не дает такие определенные предсказания, что, если они окажутся неправильными, то теория будет опровергнута. Мы принимаем теорию естественного отбора не потому, что она выдерживает попперианский тест, а потому, что она обеспечивает *правдоподобное объяснение* огромного количества данных; потому, что она оказалась плодотворной в предложении новых теорий и согласуется с развитием генетики, молекулярной биологии и т. д., и потому, что фактически предложенные альтернативные теории или были фальсифицированы, или кажутся полностью неправдоподобными в терминах фонового знания. Короче говоря, мы принимаем Дарвинову теорию эволюции путем естественного отбора как то, что Пирс называл «абдукцией», или что недавно было названо «выводом к наилучшему объяснению». Это именно тот вид вывода, которого Поппер хотел *добиться* от науки; но Попперу не слишком удастся убедить

ученых в том, что они должны отказаться от теорий, которые не являются сильно фальсифицируемыми, в тех случаях, когда эти теории обеспечивают хорошее объяснение большого количества данных, и в тех случаях, когда в этой области нет никакого правдоподобного альтернативного объяснения. В самом деле, как я отметил в другой публикации¹⁴, Поппер преувеличивает степень, до которой даже теории классической физики являются сильно фальсифицируемыми.

Мы еще далее ослабляем наше описание научного метода, допуская «вывод к лучшему объяснению» как законную форму научного вывода, даже когда выведенное «лучшее объяснение» не является сильно фальсифицируемым в смысле Поппера. Научный метод теперь стал чрезвычайно неопределенной вещью¹⁵; но мы ожидали этого, так или иначе, ввиду тех формальных результатов в индуктивной логике, которые мы описали выше. Может ли быть *настолько* неопределенное описание «научного метода» исчерпывающим? И может ли при даже таком неопределенном описании быть ясно, что ценностные суждения не подлежат проверке или не могут быть подтверждены этим методом? Ответ состоит в том, что, если научный метод описан просто как «проводить эксперименты и наблюдения так тщательно, как возможно, а затем делать выводы к лучшему объяснению и устранять теории, которые могут быть фальсифицированы решающими экспериментами», то невозможно увидеть то, что *не может* быть проверено методом, так неопределенно описанным. Предположим, например, я хочу проверить утверждение «Джон — плохой человек». Я мог бы рассуждать следующим образом: «Следующее — наблюдаемые факты: что Джон несдержан, что Джон чрезвычайно эгоистичен и что Джон — очень жестокий человек. Тот, кто несдержан, эгоистичен и жесток — *prima facie* плохой человек; поэтому Джон — плохой чело-

14. См. мою статью: Putnam H. The «corroboration» of theories // Putnam H. Mathematics, Matter, and Method. Philosophical Papers, Vol. 1. Cambridge: Cambridge University Press, 1975.

15. Альтернативным образом, мы могли бы ограничить термин «научный метод» указанием на сознательное применение принципов экспериментальной процедуры, как я рекомендовал в своей книге «Значение и моральные науки», и прекратить пытаться сделать его таким всеобъемлющим, чтобы он охватил все, что мы называем «знанием».

век». Есть два пункта, по которым против этого аргумента мог бы возразить защитник подхода, состоящего в том, что «оценочные суждения» не могут быть «научно проверены». Он мог бы возражать против последнего шага; это — шаг от *Джон жесток*, *Джон несдержан*, *Джон чрезвычайно эгоистичен* к *Джон — плохой человек*. По общему признанию это — концептуальный шаг; требование состоит в том, что есть концептуальная связь между свойствами быть жестоким, невнимательным и эгоистичным и свойством быть нравственно дурным¹⁶.

Конечно, если между моральными предикатами нет никаких концептуальных связей, то этот ход недействителен; но почему мы должны считать, что между моральными предикатами не должно быть никаких концептуальных связей? Можно спорить, что использование в аргументе ходов, описанных как «концептуальные», само по себе ненаучно; но мы, конечно, не могли бы утверждать, что в самой науке нет таких ходов. Например, если я делаю вывод из Ньютонова описания солнечной системы к утверждению, что «причина приливов — гравитационное притяжение луны», то я использую свое неформальное знание о том, что есть концептуальная связь между утверждениями о силах и утверждениями вида *А — причина В*. Слово «причина» даже не появляется в Ньютоновом описании солнечной системы и приливов; но я знаю, что гравитационная сила, которую *А* проявляет на *В*, может быть описана как причинное воздействие (массы) *А*, просто на основании понимания теории Ньютона.

16. Тот факт, что истина или вывод принадлежат к тому виду, который мы называем «концептуальным», не подразумевает, что они должны быть вполне лингвистическими по характеру (то есть истинными на основании произвольных лингвистических соглашений). Философы многих различных направлений видели, что концепции, факты и наблюдения взаимозависимы. Как мы отметили в главе 6, концепции сформированы тем, что мы наблюдаем или интуитивно постигаем, и они, в свою очередь формируют то, что мы способны наблюдать и интуитивно постигать. В этих отношениях вывод в тексте, включающем предикат «хороший», в точности аналогичен следующему выводу с предикатом «сознательный»: «Джон говорит разумно, действует осмысленно и реагирует на происходящее; поэтому Джон сознателен». Концептуальная связь здесь состоит в том, что наличие свойств «говорить разумно», «действовать осмысленно», «реагировать на происходящее» является *prima facie* причинами для приписывания сознания, именно тем способом, которым несдержанность, эгоизм и жестокость являются *prima facie* причинами для приписывания нравственной порочности.

Конечно, если мы описываем научный метод как состоящий в проведении «вывода к лучшему объяснению», или чего-то подобного, из «утверждений наблюдения», которые *сами сделаны на ценностно-нейтральном языке*, то мы можем исключить «Джон несдержан» и «Джон эгоистичен» как «утверждения наблюдения» (хотя, в некоторых конкретных случаях, было бы легче достичь соглашения по ним, чем, скажем, по вопросу о том, является ли объект *сиреневым*). Но такие утверждения постоянно имеют место в работах, например, *историков*. Маловероятно, чтобы история, клиническая психология и описание обыденного языка действительно могли избегать таких слов, как «сдержанный» и «эгоистичный». (Огромной проблемой было бы и то, где провести границу: является ли ценностно нейтральным «упрямый»? Является ли ценностно нейтральным «сердитый»? В этом отношении, является ли ценностно нейтральным суждение «зверски вывернул ей руку»?) Но в любом случае, чтобы идентифицировать *рациональность* с описанной таким образом научной рациональностью, надо считать решенным спорный вопрос познавательного статуса ценностных суждений; надо было бы сказать, что эти суждения рационально не подтверждаемы, потому что *они* – ценностные суждения, поскольку рациональность была *определена* как состоящая исключительно из сырого и нейтрального наблюдения и проведения выводов из ценностно нейтральных посылок. Но почему мы должны принять такое определение?

9. ЦЕННОСТИ, ФАКТЫ И ПОЗНАНИЕ

Я доказывал в главе 6, что «каждый факт нагружен ценностью, и каждая из наших ценностей нагружает некоторый факт». Аргумент, в двух словах, состоял в том, что *факт* (или истина) и *рациональность* — взаимозависимые понятия. Факт есть нечто, во что рационально верить, или, точнее, понятие факта (или истинное утверждение) — идеализация понятия утверждения, заслуживающего рационального доверия. «Рационально приемлемый» и «истинный» — понятия, которые зависят друг от друга. Я доказывал, что свойство быть рациональным привлекает критерии наличия *релевантности*, также как критерии рациональной приемлемости, и что все наши ценности связаны с нашими критериями релевантности. Решение, что картина мира истинна (или истинна по нашим существующим представлениям, или «столь же истинна, как и что бы то ни было») и *отвечает на уместные вопросы* (постольку, поскольку мы способны ответить на них), опирается на всю нашу систему ценностных обязательств и раскрывает ее. Не иметь ценностей значило бы также не иметь никаких фактов.

Способ, которым критерии релевантности вовлекают ценности, по крайней мере косвенно, можно увидеть, исследуя самое простое утверждение. Возьмем предложение «кошка — на коврике». Если кто-либо действительно делает это суждение в определенном контексте, то он использует концептуальные ресурсы — понятия «кошка», «на», и «коврик», — предоставляемые определенной культурой, присутствие и распространенность которых сообщают нам кое-что об интересах и ценностях этой (и почти каждой) культуры. Мы имеем категорию «кошка», потому что мы расцениваем разделение мира на *животных* и *неживотных* как существенное, и нам, далее, немаловажно, к какому *виду* данное животное принадлежит. *Релевантно*, что на этом коврике находится *кошка*, а не просто *вещь*. Мы имеем категорию «коврик», потому что мы расцениваем разделение неодушевленных вещей на *артефакты* и *не-артефакты* как существенное, и нам, далее, немаловажны *назначение* и *природа* конкрет-

ного артефакта. Важно то, что кошка находится именно на ков-
рике, а не на *чем-то*. Мы имеем категорию «на», потому что нам
важны *пространственные отношения*.

Заметим, что у нас получается: мы взяли самое банальное ут-
верждение, которое только можно вообразить: «кошка на ков-
рике», и увидели, что предположения, которые делают это ут-
верждение *релевантным* в некоторых контекстах, включают
значимость категорий *одушевленность/неодушевленность, цель и про-
странство*. Для сознания, не обладающего диспозицией расце-
нивать их как *релевантные* категории, замечание «кошка на ков-
рике» носило бы столь же *иррациональный* характер, что и
реплика «число шестиугольных объектов в этой комнате — 76»,
произнесенное в разгар встречи молодых влюбленных.

Не только самые общие факты о нашей системе ценностей
проявляются в наших категориях (*артефакты, название вида, тер-
мин пространственного отношения*), но, как мы видели в главе 6,
также и наши более определенные ценности (например, чув-
ствительность и сострадание) обнаруживаются в нашем исполь-
зовании определенных классификационных слов («сдержан-
ный», «эгоистичный»). Повторим: наши критерии релевантно-
сти опираются на всю нашу систему ценностей и раскрывают ее.

Это обсуждение проблемы релевантности имеет самое непосред-
ственное отношение к вопросу, поднятому в предыдущей гла-
ве («чем ценна рациональность?»). Если «рациональность» —
способность (или, вернее, объединенная система способнос-
тей), которая позволяет ее обладателю определить то, какие *воп-
росы* уместно задать и какие *ответы* можно уверенно принять,
то в этом и заключается ее ценность. Но не требуется никаких
аргументов, чтобы показать, что *такая* концепция рациональ-
ности так же ценностно нагружена, как и само понятие релеван-
тности.

Можно возразить, что я смешал в одну кучу факторы, кото-
рые на самом деле не связаны друг с другом, и тем самым совер-
шил своего рода подлог. Сам тот факт, что я говорил о двух фак-
торах, рациональной приемлемости и релевантности, свиде-
тельствует, как можно утверждать, об устойчивости и постоян-
стве соответствующей дихотомии, подобной дихотомии факта/
ценности. Согласно концепции, которую возражающий имеет
в виду, рациональный человек — это тот, кто может провести

различие между тем, что является, а что не является обоснованно утверждаемым. Возражение состояло бы в следующем: то, что человек считает *интересным* или *важным*, или *релевантным*, может иметь отношение к оценке его характера или даже его умственного здоровья, но не к его *познавательной* рациональности.

Приемлемость и релевантность взаимозависимы, однако, в любом реальном контексте. Использование любого слова — будь то слово «хороший» или «сознательный», или «красный», или «магнитный» — отсылает нас к истории, традиции наблюдения, обобщения, практики и теории.

Это также вовлекает нас в деятельность по *интерпретации* этой традиции и ее приспособлению к новым контекстам, ее расширению и критике. Можно интерпретировать традиции по-разному, но вообще невозможно применять слово, если мы находимся полностью вне традиции, которой оно принадлежит. А нахождение внутри традиции, конечно, влияет на то, что считать «рационально приемлемым». Если бы можно было использовать единый метод, чтобы проверить любое утверждение вообще, независимо от того, какие понятия оно содержит, то предложенное отделение способности проверять утверждения от владения *релевантным* множеством понятий могло бы быть логичным. Однако мы уже видели, что нет никакой причины соглашаться с мифом о существовании одного единственно правильного метода.

ТЕОРИЯ ДВУХ КОМПОНЕНТОВ

Существующие у нас интуитивные представления о рациональности, по всей видимости, находятся в конфликте; конечно, ни одна философская теория не выглядит примиряющей их все. С одной стороны, просто неверно, что мы *никогда* не судим о целях как о рациональных или иррациональных; с другой стороны, когда мы сталкиваемся со случаем, подобным случаю гипотетического «рационального нациста», мы не видим, как оправдать критику такой разумно разработанной и рассматриваемой системы целей как *иррациональной*, даже если мы находим ее нравственно отталкивающей.

Один из способов, которым, как было предложено, мы могли бы решить эти проблемы, состоит в следующем: принимая

резкую дихотомию факта/ценности, мы можем обосновать характеристику человека, интересующегося лишь числом волос на головах людей, как иррационального на том основании, что его восприятие *фактов* неадекватно (что здесь означает «адекватный», конечно, остается проблемой). Нацист лишь не соглашается с нами в споре о *ценностях*, вот почему он — не иррационален. Промежуточные случаи могут быть рассмотрены, скажем, в перспективе, предложенной Бернардом Уильямсом. Тогда, в частности, наш аргумент против фетишизма метода, состоявший в частности в том, что мы не можем, не попадая в замкнутый круг в доказательстве, устранить из «данных наблюдения» такие описательные суждения, как «Джон сдержан», может встретить то возражение, что морально-описательный словарь обыденного языка имеет одновременно два «компонента значения». Один из компонентов — *фактический*; есть некоторые общепринятые *стандарты* сдержанности, и суждение «Джон сдержан» передает информацию, что Джон соответствует этим стандартам¹. Но есть также *эмоциональный* компонент значения: суждение «Джон сдержан» передает «пред-отношение» к некоторому аспекту поведения Джона. Утверждение состоит в том, что рациональным признается принятие *фактического* компонента утверждения «Джон сдержан»; принятие эмоционального компонента значения, разделяющего «пред-отношение», не является ни рациональным, ни иррациональным.

Понятие «фактического» здесь происходит от взглядов, которых данный философ придерживается относительно Содержания Мира. Для философа-материалиста «фактический» компонент в значении любого утверждения должен состоять в утверждении, выразимом в слове *физической науки*. Но здесь

1. Однако, остается проблемой, что здесь означает выражение «стандарты». Если требование, что «Джон сдержан», является описательно истинным (то есть «фактический компонент» является истинным), если бы и только если *большинство говорящих согласилось бы, что Джон сдержан*, то из этого анализа «фактического компонента» следовало бы, что нет такого человека, кого восприимчивый судья *правильно* классифицировал бы как «сдержанного» даже при том, что *большинство* говорящих *не соглашается* с этим. Такая теория «фактического компонента» просто составила бы требование, что все истины (по крайней мере, о «стандартах») должны быть «публичными», но почему мы должны верить *этому*, если только мы не занимаем мажоритарную позицию? (Как мы видели в главе 5, утверждение, что *всякая* истина публична, опровергает себя.)

сразу возникает трудность, сопоставимая с трудностью, с которой столкнулся феноменализм.

Феноменализм, как мы помним, считает, что все значащие утверждения переводимы без остатка в утверждения об ощущениях. «Физикализм» (как стали называть тот тип материализма, который мы обсуждаем) считает, что все «фактическое» значение любого утверждения может быть переведено без остатка на язык физики. И, повторим, эта доктрина представляется ложной.

Чтобы увидеть, почему это так, рассмотрим не морально-описательное утверждение, а психологическое, скажем, «Х думает о Вене». Очевидно, что даже если для того, чтобы думать о Вене, есть необходимые и достаточные условия, выражимые в терминах состояний мозга, или чего бы то ни было, произвольного организма, то потребуются невообразимо совершенная неврологическая теория (или, возможно, теория функционалистской психологии), чтобы сказать, каковы они. Условия истинности этого утверждения зависят от контекста, они относительно к интересам и неопределенны. Нет никакой причины думать, что даже «в принципе» существует *конечное* выражение на языке физической теории, которое (в любом физически возможном мире) является истинным об Х, если и только если Х думает о Вене. Мало того, что может быть *ложно*, что существует конечный эквивалент на физикалистском языке для утверждения обыденного языка «Х думает о Вене»; даже если такой эквивалент существует, эта эквивалентность была бы эквивалентностью на основе эмпирической теории или группы теорий, которые сами не известны (возможно, они так сложны, что люди *никогда* не будут *знать* их), и которые наверняка не являются частью *значения* «Х думает о Вене». Короче говоря, ложно, что «Х думает о Вене» *означает* «Х находится в таком-то и таком-то (физически или функционально указанном) состоянии мозга».

Что справедливо для «Х думает о Вене», будет справедливо и для любого предиката обыденного языка, условия применения которого не слишком хорошо согласуются с теми, которые управляют физическими понятиями. «Х сдержан», и даже «Х — коричневый», «Х — землетрясение» и «Х — человек» — также не *переводимо* на язык «физической теории». Это означает, что если *есть* два компонента значения «Х сдержан», то единственное

описание, которое мы можем дать «фактическому значению» утверждения — это то, что оно истинно, если и только если *X* *сдержан*. А это делает понятие «фактического компонента» тривиальным.

Крах теории «двух компонентов» не ведет к отрицанию того, что «*X* сдержан» обычно имеет некоторую эмоциональную силу. Но так происходит не всегда. Как мы указали в главе 6, мы можем использовать утверждение «*X* сдержан» для многих целей: чтобы оценивать, описывать, объяснять, предсказывать и так далее. Различение способов *использования* этого утверждения не требует, чтобы мы отрицали существование такого *утверждения*, как «*X* сдержан».

МУР И «НАТУРАЛИСТИЧЕСКАЯ ОШИБКА»

Утверждение Вебера о том, что «ценностные суждения» не могут быть рационально подтверждены, стало источником существующей дихотомии факта/ценности; однако помимо всего прочего, эта дихотомия была укреплена Дж. Е. Муром (вопреки его собственным намерениям). В те времена, когда Бертран Рассел и Джон Мейнард Кейнс, наряду с другими будущими членами Блумсберийской группы, были еще молодыми студентами, Мур уже приводил доводы в пользу тезиса о том, что Добродетель — «не-природное» свойство. Иными словами, он выступал в защиту тезиса, находящегося полностью вне физикалистской онтологии естествознания. Его защита не-натурализма имела печальную историю; его студенты, возможно, были убеждены Муром, что есть такие вещи, как «не-природные» свойства (хотя Рассел, по крайней мере, позже утратил эту веру), но более поздние философы натуралистического толка склонны были полагать, что Мур осуществил *reductio ad absurdum*² идеи, что есть такие вещи как свойства ценности. В 1930-х годах Чарльз Стивенсон и логические позитивисты выдвинули «эмоциональную теорию этики», согласно которой «*X* является хорошим» значит «я одобряю *X*, поступайте так же!», или что-то в этом роде. Ценностные *свойства* отклоняются на эпистемологических основаниях, но даже еще в большей степени — на онтологических ос-

2. Сведение к абсурду (лат.). — Прим. пер.

нованиях; как это недавно выразил Джон Маккай³, такие вещи, как *ценностные отношения*, совместимы с естествознанием, но такие вещи, как *ценностные свойства*, не совместимы с естествознанием. Ценностные свойства, как утверждает Маккай, являются «онтологически странными», то есть они — забавные таинственные свойства, в существование которых научно просвещенные люди не должны продолжать верить.

Аргумент Мура о том, что Добродетель не может быть физическим свойством («природным» свойством), состоял в том, что если «Добродетель» является тем же свойством, что и «способствовать максимизации совокупной полезности» (или еще каким-то природным свойством, физическим или функциональным), то

- (1) «это действие не является добродетельным даже при том, что оно способствует максимизации совокупной полезности»,

является внутренне противоречивым утверждением (а не только ложным).

Но даже утилитарист не стал бы утверждать, что (1) *внутренне противоречиво*. И это показывает, как требует Мур, что, хотя свойства «быть добродетельным» и «являться способствующим максимизации совокупной полезности» могут быть *коррелирующими* свойствами, они не могут быть *тождественными*.

Однако аргумент Мура опирается на такие исходные допущения, которые я и многие другие философы языка сегодня отклонили бы. Прежде всего, он неявно отрицает, что возможна такая вещь как *синтетическая идентичность свойств*. Но, как я указал в главе 4, это исключит такие принятые научные открытия, как открытие, что величина *температуры* является той же самой величиной, что и величина *средней молекулярной кинетической энергии*. (Можно использовать «доказательство» Мура, чтобы показать, что температура должна быть, фактически, «не-природным свойством». Мы не *противоречим себе*, когда говорим «*x* имеет температуру *T*, но *x* не имеет средней молекулярной кинетической энергии *E*», где *E* — значение кинетической энергии, которая соответствует температуре *T*, даже если это утверждение

3. Mackie J. Ethics, Inventing Right and Wrong. Penguin, 1977.

всегда ложно как эмпирический факт. Таким образом, должен был бы заключить Мур, температура лишь *коррелирует* со средней молекулярной кинетической энергией; эти два свойства не могут быть буквально *идентичны*). Фактически, Мур объединяет *свойства* и *понятия*. *Есть* такая идея свойства, в которой тот факт, что два понятия являются различными (скажем, «температура» и «средняя молекулярная кинетическая энергия»), несколько не улаживает вопрос, являются ли соответствующие *свойства* различными. (А открытие того, сколько существует фундаментальных физических свойств, обнаруживает не нечто о *понятиях*, но нечто о *мире*). *Понятие* «добродетели» не может быть синонимичным с каким-либо физикалистским понятием (в конце концов, морально-описательный язык и физикалистский язык представляют собой чрезвычайно различные «версии», в смысле Гудмена), но из этого не следует, что *быть добродетельным* — не то же свойство, что *быть Р*, для некоторого подходящего физикалистского (или, лучше, функционалистского) *Р*. В общем, можно сказать, что наглядно выученный термин для свойства (например «имеет высокую температуру») не синонимичен с теоретическим определением этого свойства; чтобы выяснить, чем является температура (и, как могли бы предложить некоторые философы, чем является *добродетель*), требуется эмпирическое и теоретическое исследование, а не только лингвистический анализ, не только рефлексия над значениями.

Через несколько лет после того, как я представил понятие «синтетической идентичности свойств», в философию языка вошла идея, которая расширяет и проясняет сделанные мной замечания — это идея Сола Крипке о «метафизически необходимых» истинах, которые должны быть изучены *опытным путем*, «эпистемически случайных необходимых истинах»⁴. Наблюдение Крипке, в применении к случаю температуры/кинетической энергии, состоит в том, что, если кто-то описывает логически возможный мир, в котором люди имеют ощущения тепла и холода, в котором есть горячие на ощупь и холодные на ощупь объекты и в котором эти ощущения тепла и холода *объясняются иным механизмом, чем средняя молекулярная кинетичес-*

4. Kripke S. Naming and Necessity. Harvard, 1980. (Лекции, прочитанные в Принстоне в 1970 году).

кая энергия, то мы не говорим, что он описал возможный мир, в котором *температура – не средняя молекулярная энергия*. Скорее мы скажем, что он описал мир, в котором *некоторые механизмы, иные, чем температура, делают некоторые объекты горячими и холодными на ощупь*. Как только мы принимаем «синтетическое утверждение идентичности», что температура является средней молекулярной кинетической энергией (в *действительном* мире), ничто не считается возможным миром, в котором температура не является средней молекулярной кинетической энергией.

Утверждение, которое является истинным в каждом возможном мире, традиционно называется «необходимым». Свойство, которым что-либо обладает в каждом возможном мире, традиционно называется «сущностным». В этой традиционной терминологии, Крипке говорит, что истина «температура – средняя молекулярная кинетическая энергия» является *необходимой* истиной *даже при том, что мы не можем знать это априорно*. Это утверждение является эмпирическим, но необходимым. Иными словами, быть средней молекулярной кинетической энергией – необходимое свойство температуры. Мы обнаружили *сущность* температуры при помощи *эмпирического* исследования. Эти идеи Крипке имели широкое влияние на философию языка, метафизику и философию математики; они носили крайне разрушительный для аргумента Мура характер. Мур отталкивался от факта, что (1) *может лишь быть случайно ложно*, что быть P (для некоторого подходящего природного свойства P) не может быть необходимым свойством добродетели; именно это отрицает новая теория необходимости. Все, что можно законно вывести из факта, что (1) не является внутренне противоречивым – это то, что «добродетельный» не синонимично с «способствующий максимизации полезности» (не синонимично с P , для любого термина P в физикалистской версии мира). Из этого факта отсутствия синонимии *слов* ничего не следует отсутствия идентичности *свойств*, равно как ничего не следует о сущности добродетели.

Рут Анна Патнэм указала, что не работает и другой общий аргумент, при помощи которого пытаются доказать, что совершенство не может быть природным свойством⁵. Это – аргумент,

5. Putnam R. A. Remarks on Wittgenstein's Lecture on Ethics // Language, Logic, and Philosophy. Proceedings of the 4th Intern. Wittgenstein Symposium / Ed. Haller et al. Vienna, 1980.

что «*Х является добродетельным*» имеет «эмоциональную силу», «выражает пред-отношение» и т. д.

Как указывает Рут Анна Патнэм, слабость этого аргумента состоит в том, что многие описательные предикаты естественным путем *приобретают* эмоциональную силу. В нашей культуре «размазывает еду по рубашке» имеет большую отрицательную эмоциональную силу, хотя эта фраза буквально представляет собой *описание*. Любое слово, означающее нечто, что люди наделяют *ценностью* в культуре (или чему приписывают негативную ценность), будет склонно приобретать эмоциональную силу.

Слово «добро», в его моральном смысле, применяется ко многим вещам. Некоторые из них — добродетельные состояния сознания, например — могут получить природную ценность, быть естественно оценены: частью содержания самого этого состояния сознания может быть то, что человек ценит нахождение в этом состоянии. Я не хочу предлагать обратное: что любое состояние сознания, которое естественно оценено в этом смысле, добродетельно; это было бы явно ложно. Предположим, что «добро» определено так, что те вещи, которые естественно оценены и найдены таковыми, что нет причин приписывать им негативную ценность или отказывать им в ценности (как тому есть причина для некоторых естественно ценных, но вызванных наркотиками состояний), считаются «добром».

Тогда можно ожидать, что утверждение, что нечто добродетельно, будет иметь положительную «эмоциональную силу» *в силу природы этого свойства*. Даже если не принимать позицию консеквенциализма (то есть считать, что *все*, у чего достаточно добродетельные *последствия*, само добродетельно), нет сомнения, что наиболее общая причина того, чтобы называть действие добродетельным, состоит в том, что это действие имеет добродетельные последствия, среди которых может быть то, что оно максимально способствует состояниям или ситуациям, наделенным естественной ценностью; таким образом, опять сама природа свойства объясняет, откуда у описания «пред-эмоциональная сила».

Маккай защищает свое утверждение, что добродетель является онтологически «странным» свойством, вводя в качестве посылки допущение о том, что нельзя *знать*, что нечто является добродетельным, *не находясь* «предварительно» в определен-

ном отношении к этому нечто. Это ведет к использованию эмотивизма для доказательства эмотивизма. Часто изображают, как черти в аду используют понятие «добродетель» с *отрицательной* эмоциональной силой (они могут сказать о ком-то: «У него при- скорбная склонность к моральной добродетели»); в отличие от Маккая, я не нахожу, что такое использование будет лингвистически неподходящим или вовлекающим в какое бы то ни было противоречие. И разве мы не встречаемся с тем, как люди говорят: «Я знаю, что делать такие вещи дурно, но мне все равно»? Как указала Филиппа Фут, можно даже отказаться от обращения к этике, заявляя: «Я не собираюсь быть добродетельным»⁶. Возможность таких высказываний показывает, что в то время как действительно есть различие между *описательным* использованием языка и *предписательным* или *рекомендательным* использованием, это различие в использованиях — не просто функция словаря. «Описательные» слова могут использоваться, чтобы хвалить или обвинять («Он размазывает еду по галстуку»), а «оценочные» слова могут использоваться, чтобы описывать и объяснять. (Рассмотрим следующий диалог: «Джон, должно быть, был исключительно хорошим человеком, раз так поступил». «Нет, он никогда не был образцом морали, если не наоборот; но у него, наверное, была способность к самопожертвованию, о которой мы никогда не подозревали». Здесь *моральный* язык используется с *объяснительной* функцией). Вновь повторяя замечание Рут Анны Патнэм, скажем, что из того факта, что «быть добродетельным», используется для рекомендаций, никак не следует, что добродетель не является свойством.

Профессор Патнэм указывает, что в аргументе Маккая есть, однако, нечто *правильное*. Некоторые моральные выражения бесспорно содержат встроенную ориентацию к действию. «Должен», «обязан» и «правильный» дают явные примеры таких «содержащих руководство к действию» слов. Как она указывает, проблема «есть/должен» не совпадает с проблемой «факт/ценность». «Я не собираюсь делать то, что должен» звучит намного более странно, чем «я не собираюсь быть добродетельным» (а «я не собираюсь делать то, что обязан» звучит безумно).

6. См., например: *Foot P. Virtues and Vices and Other Essays in Moral Philosophy*. Berkeley: University of California Press; Oxford: Blackwell, 1978; *Moral Relativism*. Lindley Lecture, 1978. Lawrence: University of Kansas Press, 1979. — *Прим. пер.*

Маккай указывает, что ни у какого физического свойства не может быть встроенной связи с действием (или с одобрением действия), и заключает, что такие предикаты, как «быть правильным поступком» и т. д. являются «онтологически странными». Но (помимо того, что он зависит от предположения, что физикалистская версия мира является Единой Истинной Теорией), этот аргумент доказывает слишком многое. Ведь некоторые *эпистемические* предикаты (например, «рационально приемлемый», «обоснованное убеждение») также содержат руководство к действию (рассматривая «действие» в широком смысле, так, чтобы принятие утверждения считалось действием). Можно говорить «Х — хороший поступок» и «Весьма очевидно, что Y» и не стремиться к выполнению (или предписанию) X или принятию Y; но если мы говорим «Х — правильное действие для выполнения в этой ситуации» или «Полагать, что Y, полностью обоснованно», то мы ориентированы на выполнение (или предписание) X и на принятие Y. «Обоснованный» (в случае убеждений) имеет характеристику быть руководством к действию настолько же, насколько она есть у предиката «правильный» в моральном смысле.

Если мы теперь воспроизведем аргумент Маккай и придем к выводу, что нет такого свойства, как *быть обоснованным*, но есть лишь «отношение обоснования», то мы придем к полному релятивизму. Перед тем, как зайти так катастрофически далеко, мы должны сделать паузу, чтобы увидеть, *почему* предикаты руководства к действию кажутся «онтологически странными» убежденному физикалисту.

В главе 2 я доказывал, что сама *референция* должна казаться «онтологически странной» убежденному физикалисту. Если действительно существуют только физикалистские свойства и отношения, то референция, чтобы вообще существовать, должна быть физикалистским отношением — но тогда проблема оказывается, как мы видели, в огромном изобилии «кандидатов». Есть бесконечное число допустимых отношений референции (и все они — физикалистские, или, по крайней мере, натуралистические, если мы считаем *теорию множеств* частью натуралистической версии мира). Если бы одно из них было *единственным* отношением референции, то сам *этот* факт был бы *элементарным* метафизическим фактом очень странного вида.

Такой факт делало бы странным то, что мы вкладываем в само наше понятие Природы некоторый *нейтралитет*, некоторую *бессознательность*. У Природы, как предполагается, нет никаких интересов, намерений или точки зрения. Если это правильно, то как *могло бы* быть метафизически выделено одно допустимое отношение референции?

Именно эта бессознательность Природы заставляет предикаты руководства к действию «быть правильным» и «быть обоснованным убеждением» казаться «странными». Если бы некоторое физикалистское свойство Р было *идентично* с моральной справедливостью или с эпистемологическим обоснованием, это *было бы* «странным» — странным по точно той же самой причине, что было бы «странно», если бы *референция* была физикалистским отношением. Это было бы так, как если бы у самой природы наличествовали ценности, в моральном случае, или референциальные намерения, в семантическом случае.

Поэтому я думаю, что Мур был прав (даже если его аргументы не приемлемы), утверждая, что «добродетельный», «правильный» (как и «обоснованное убеждение», «указывает на» и «истинный») не идентичны с физикалистскими свойствами и отношениями. *Это* показывает не то, что добродетель, справедливость, эпистемическое обоснование, референция и истина не существуют, но что монистический натурализм (или «физикализм») является неадекватной философией.

СНОВА О «РАЦИОНАЛЬНОМ НАЦИСТЕ»

Ранее нас беспокоило то, что мы не видели, как показать, что некий гипотетический «совершенно рациональный нацист» имел иррациональные цели. Возможно, проблема в том, что мы слишком просто идентифицировали вопрос о рациональности *нациста* (как кого-то, кто имеет мировоззрение или мировоззрения) с вопросом о рациональности *целей* нациста. Если «у» нациста нет никакой цели, к которой мы можем обратиться, то кажется странным диагностировать ситуацию, говоря «Карл имеет иррациональные цели». Даже если это составит часть того заключения, к которому мы пришли в итоге, *первое*, что мы скажем, — то, что Карл имеет *чудовищные* цели, а не иррациональные.

Но наш вопрос здесь, если мы собираемся обсуждать рациональность Карла вообще, стоит о нелогичности его *убеждений и аргументов*, а не его цели.

Предположим сначала, что Карл заявляет, что нацистские цели нравственно *правильны и хороши* (как вообще и поступали нацисты на самом деле, а не в примерах философов). Тогда, *фактически*, он будет говорить ерунду. Он будет выносить всевозможные ложные «фактические» суждения, например, что демократические государства управляются «еврейским заговором», и выдвинет моральные суждения (например, что быть «арийцем» значит быть *обязанным* поработить неарийские расы «господствующей расой»), для которых у него нет хороших аргументов. Понятие «хорошего аргумента», к которому я обращаюсь, является *внутренним* по отношению к обычному моральному дискурсу; но это — уместное понятие, если нацист пытается оправдать себя *в пределах* обычного морального дискурса.

Предположим, с другой стороны, что наш нацист *не признает* обычные моральные понятия в целом (как наши гипотетические Супер-Бентамиты). Я доказывал, что культура, которая не признает обычные моральные понятия или заменила их понятиями, полученными из другой идеологии и моральной перспективы, утратит способность адекватно и ясно *описывать* обычные межличностные отношения, социальные и политические события *в наших существующих представлениях*. Конечно, если эти другие идеология и моральная перспектива *превосходят* нашу существующую моральную систему, то эта замена может быть хороша и мудра; но если другая идеология и моральная перспектива *плохи*, особенно если они исковерканы и чудовищны, то результатом будет просто неадекватное, неясное, отталкивающее представление межличностных и социальных фактов. Конечно, выражения «неадекватный, неясный, отталкивающий» отражают ценностные суждения; но я доказывал, что выбор концептуальной схемы *обязательно* отражает ценностные суждения, и выбор концептуальной схемы — это все, что затрагивает *познавательная* рациональность.

Даже если индивидуальный нацист не теряет способности использовать наш существующий моральный описательный словарь, даже если он сохраняет в голове старые понятия (так некоторые ученые, возможно, все еще знакомы со средневековым

понятием «галантности» и способны использовать его), все последние (существующие у нас моральные описательные понятия вроде «сдержанный», «сострадательный», «справедливый», «честный») не будут теми понятиями, которые он использует в жизни: в действительности они не будут фигурировать в его строении мира.

Я хочу еще раз подчеркнуть, что я не говорю, что нацистом быть *плохо* потому, что это ведет к деформированным и иррациональным *убеждениям*. Нацистом быть плохо потому, что это ведет к определенным *поступкам*. Нацист зол, *а также* у него иррациональное мировоззрение. Эти два факта взаимосвязаны; но это не подразумевает, что нацист зол прежде всего *потому, что* у него иррациональное мировоззрение в смысле, что иррациональность его мировоззрения *составляет* зло. Однако, как мне кажется, *есть* смысл, в котором мы можем говорить здесь о рациональных или иррациональных *целях* такие цели, что, если мы принимаем и преследуем их, то мы придем или к предложению безумных и ложных аргументов в их пользу (если мы ставим задачу их обоснования в пределах нашей нормальной концептуальной схемы), или к принятию альтернативной — иррациональной — схемы представления обычных морально-описательных фактов (например, того, что кто-либо сострадателен) — такие цели нужно назвать «иррациональными». В конце концов между использованием рациональной концептуальной схемы в описании и восприятии нравственно релевантных фактов и наличием некоторых общих типов целей в противоположность другим существует определенная связь.

«Но что, если нацист не приводит никакой причины для того, чтобы быть нацистом кроме «мне нравится так действовать?» Это естественный вопрос, но, конечно, естественный ответ на него и есть правильный ответ: в таком случае поведение нациста, кроме того, что является дурным, было бы также *полностью произвольно*. Заметим, что «произвольный» — одно из слов, которые я назвал «морально-описательными», то есть слово, которое может использоваться, без изменения обозначения, для оценки (в этом случае для обвинения), для описания («Джон весьма произвольно решил сменить работу»), объяснения (или указания, что никакое объяснение некоторого вида не может быть дано) и т. д. Действительно, когда я только сказал, что ре-

шение Карла быть нацистом (в описанном случае) будет полностью произвольно, я прежде всего *описывал*, а не оценивал. Многое из того, что я делаю, весьма буквально, произвольно — например, выбор одной дорожки через университетский городок, а не другой; но это не подразумевает, что с этими действиями что-то не так. (И предметы просто слишком тривиальны). Даже если я делаю нечто важное «произвольно» — скажем, меняю работу, — если у меня нет семейных обязанностей и т. д., то это может быть просто мое *право*. Но если это такое действие, которое *требует обоснования*, то, выполняя его произвольно и без обоснования, я подвергнусь законному осуждению. Принятие решения, которое неблагоприятно затрагивает жизни других (и, возможно, собственную жизнь) в значительной степени без обоснования, как произвольный и преднамеренный (еще одно из этих моральных описательных слов!) акт, является парадигматическим примером иррациональности, и не только иррациональности, но и извращенности.

Мы начали наше обсуждение в главе 7, с рассмотрения требования Бентама, что, «если отбросить предрассудки», игра в «блошки» столь же хороша, как «искусства и науки музыки и поэзии». Согласно Бентаму, единственной причиной, почему поэзия лучше, чем игра в «блошки», в конечном счете является тот грубый факт, что поэзия дает большее удовлетворение, чем игра в «блошки» (или дает удовлетворение большему количеству людей, или и то, и другое). Две основных неправильности этого подхода таковы: одна состоит в том, что само по себе «удовлетворение» (или «личный интерес») не может быть целью никого, кто не имел бы *других* целей. Если бы у меня не было никакой *другой* цели, кроме «моего благосостояния», то мое «благосостояние» было бы бессмысленным понятием — замечание, которым мы обязаны епископу Батлеру⁷. Что более важно, некоторые варианты *удовлетворения* лучше и «благороднее» других, и можно назвать причины того, почему это так. Поэзия и музыка дают утешение, повышают нашу чувствительность, дают важные способы самовыражения многим людям, включая многих из наиболее одаренных людей, которых произвела человеческая раса. Определение этих причин для оценки некоторых ва-

7. Сэмюэл Батлер (1835 — 1902) — английский писатель. — Прим. пер.

риантов удовлетворения выше других как «предубеждений» действительно близко связано и с «теорией двух компонентов», и с идеей, что ценностные свойства «онтологически странны». Бен-там работает с моделью «нейтральных фактов» и «произвольных предубеждений». В самом деле, назвать предпочтение поэзии «предубеждением» — лишь способ, которым Бен-там предлагает, что тот факт, что поэзия дает большее удовлетворение, чем игра в «блошки», является *единственным* не «произвольным» соображением в их сравнении; любое предпочтение одного *вида* удовлетворения другому (как он предлагает) произвольно. Но это просто ложно, учитывая действительное место понятия «произвольного» предпочтения в нашей концептуальной схеме, и бессмысленно, если термин «произвольный» вынут из схемы, которой он принадлежит. (Точно так же утверждение, что предпочтение поэзии игре в «блошки» — предубеждение, *буквально ложно*). Предлагается, что так или иначе онтологически законно признать, что есть такие вещи как (варианты или частные случаи) *удовлетворения*, но онтологически незаконно признать такие вещи, как повышенная чувствительность, расширенные диапазоны значения и метафоры, способы выражения и самореализации и так далее. Идея о том, что ценности — не часть Содержания Мира, и идея о том, что «ценностные суждения» выражают «предубеждение», являются двумя сторонами одной и той же медали.

Мы исследовали вопрос, могут ли «ценностные суждения» быть рационально поддержаны. Мы видели, что различные отрицательные ответы опираются на сомнительные философские предположения: что сама рациональность хороша только для «предсказания» или только для достижения «согласия», или что есть только один Метод получения истины (где иногда единственными критериями «истины» считаются предсказание и согласие), или что ценностные суждения имеют «два компонента значения», или что ценностные свойства «онтологически странны». Позиция, которую я защищал, состоит в том, что любой выбор концептуальной схемы предполагает ценности, а выбор схемы описания обычных межличностных отношений и социальных фактов, не говоря уже о размышлении о собственном жизненном плане, привлекает, среди прочего, *моральные* ценности. Нельзя выбрать схему, которая просто «копирует»

факты, потому что *никакая* концептуальная схема не является простой «копией» мира. Само понятие истины зависит по содержанию от наших стандартов рациональной приемлемости, а они, в свою очередь, опираются на наши ценности и предполагают их. Выражаясь схематично и слишком кратко, я скажу, что теория истины предполагает теорию рациональности, которая в свою очередь предполагает нашу теорию пользы.

«Теория добродетели», однако, не только носит программный характер, но и сама зависит от предположений о человеческой природе, об обществе, о вселенной (включая теологические и метафизические предположения). Мы должны пересматривать нашу теорию добродетели (какова она есть) снова и снова, поскольку наше знание увеличивается и наше мировоззрение изменяется.

Стало ясно, что в концепции, которую я защищаю, вообще нет такой вещи, как «основание». Здесь люди начинают волноваться: не приблизились ли мы к подходу, что нет никакого различия между «обоснованным» и «обоснованным согласно нашим представлениям» (релятивизм) или даже «обоснованным согласно *моим* представлениям» (разновидность солипсизма)?

Мы действительно придем к позиции солипсизма, если попытаемся оставаться вне концептуальной системы, которой принадлежит наша концепция рациональности, и одновременно будем претендовать на то, что предлагаем более «рациональное» понятие рациональности! (Многие мыслители впали в ошибку Ницше, сообщая нам, что у них есть «лучшая» этика, чем вся традиция; в каждом случае они лишь породили чудовище, потому что все, что они *смогли* сделать, — это *произвольно* вырвать некоторые ценности из контекста при игнорировании других). Мы можем надеяться дать более рациональную *концепцию* рациональности или лучшую *концепцию* этики, только если мы работаем *внутри* нашей традиции (с ее отзвуками греческой агоры, Ньютона и так далее, в случае рациональности, и с ее отзвуками Священного Писания, философов, демократических революций и так далее, в случае этики); но это вовсе не значит, что теперь с нашими концепциями все вполне разумно и хорошо. Мы не заманены в индивидуальный солипсистский ад, как в ловушку, но приглашены участвовать в подлинном человеческом диалоге, сочетающем общность с индивидуальной ответственностью.

Есть ли в идеале у этого диалога конечная остановка? Есть ли *истинная* концепция рациональности, *истинная* этика, даже если все, что у нас когда-либо имелось — это наши *концепции* их? Здесь философы разделяются, подобно всем остальным. Ричард Рорти в своем Президентском обращении⁸ к Американской Философской Ассоциации настоятельно высказался в пользу подхода, что есть только диалог; никакой идеальный конец не может быть установлен и не должен рассматриваться как необходимый. Но чем утверждение, что «есть только диалог», отличается от самоопровергающего релятивизма, который мы обсуждали в главе 5? Сам факт, что мы говорим о наших различных концепциях как о различных концепциях *рациональности*, устанавливает *Grenzbegriff*, предельное понятие идеальной истины.

8. Rorty R. Pragmatism, Relativism and Irrationalism // Proceedings and Addresses of the American Philosophical Association. August 1980. См. также: Rorty R. Philosophy and the Mirror of Nature. Princeton University Press, 1979. [Ср. русский перевод: Рорти Р. Философия и зеркало природы / Пер. с англ. В. В. Целищева. Новосибирск: Изд-во НГУ, 1997. — *Прим. пер.*].

ПРИЛОЖЕНИЕ

Вот теорема, упомянутая в главе 2.

Теорема: Пусть L — язык с предикатами $F_1, F_2, \dots, F_k$ (не обязательно монадические). Пусть I — интерпретация, в смысле назначения интенционала каждому предикату L . Тогда, если I нетривиально в том смысле, что по крайней мере один предикат имеет экстенционал, который не является ни пустым, ни универсальным по крайней мере в одном возможном мире, то существует вторая интерпретация J , которая не согласуется с интерпретацией I , но которая делает истинными в каждом возможном мире те же самые предложения, что и I .

Доказательство: Пусть $W_1, W_2, \dots$ — возможные миры в некотором упорядочении, и пусть U_i — множество возможных индивидов, которые существуют в мире W_i . Пусть R_{ij} — множество, которое является экстенсионалом предиката F_i в возможном мире W_j согласно I (если F_{ij} не монадично, то R_{ij} будет множеством n_i членов, где n_i — число мест аргумента от F_i). Структура $\{U_i; R_{ij} (i = 1, 2, \dots, k)\}$ является «интендированной моделью» L в мире W_j относительно I (то есть U_j является вселенной дискурса L в мире W_j и (для $i = 1, 2, \dots, k$) R_{ij} — экстенционал предиката F_i в W_j).

Если по крайней мере один предикат, скажем, F_u , имеет экстенционал R_{uj} , который не является ни пустым, ни всеми экстенсионалами U_j , выберем такую перестановку P_j для U_j , что $P_j(R_{uj}) \neq R_{uj}$. Иначе, пусть P_j будет тождеством. Поскольку P_j является перестановкой, структура $\{U_i; P_j(R_{ij}) (i = 1, 2, \dots, k)\}$ изоморфна $\{U_i; R_{ij} (i = 1, 2, \dots, k)\}$, и такова же модель для тех же самых предложений L (то есть для предложений L , которые являются истинными при I в W_j).

Пусть J — такая интерпретация L , которая назначает предикату F_i ($i=1, 2, \dots, k$) следующий интенционал: функция $f_i(W)$, значением которой в любом возможном мире W_j является $P_j(R_{ij})$. Другими словами, экстенционал F_i в каждом W_j при интерпретации J определен как $P_j(R_{ij})$. Поскольку $\{U_i; P_j(R_{ij}) (i = 1, 2, \dots, k)\}$ является моделью того же самого множества предложений, что и $\{U_i; R_{ij} (i = 1, 2, \dots, k)\}$ (изоморфизмом), те же самые предложе-

ния истинны в каждом возможном мире при I и при J , и J отличается от I в каждом мире, в котором по крайней мере один предикат имеет нетривиальный экстенционал, что и требовалось доказать.

Комментарий: Если в данном мире W_j есть два непересекающихся множества, которые являются экстенционалами предикатов L в W_j при I — скажем, множество кошек и множество собак — то, если собак больше, чем кошек (соответственно, по крайней мере столько же кошек, сколько собак), мы можем взять любое множество собак той же величины, что и множество кошек (соответственно, любое множество кошек той же величины, что и множество собак) и выбрать такое P_j , которое проецирует выбранное множество собак на множество кошек (соответственно, выбранное множество кошек на множество собак) и наоборот; это будет гарантировать, что при J экстенционал первого предиката (того, экстенционал которого при I — множество кошек) является множеством собак при J в W_j или экстенционале второго предиката (того, экстенционал которого при I — множество собак) является множеством кошек при J в W_j .

Второй комментарий: Если есть объекты — скажем, «ощущения» — которые мы не хотим переставлять, потому что рассматриваем предикаты этих объектов как «абсолютные» в некотором смысле, мы просто постулируем, что перестановки P_j должны быть тождеством этих объектов. Это будет накладывать одни и те же ограничения на любой предикат L для этих привилегированных объектов при I и при J в каждом мире.

Третий комментарий: Поскольку предложения получают логически эквивалентные условия истинности при I и при J , из этого следует, что, по стандартной «семантике возможных миров», контрфактуальные кондиционалы также сохраняются.

«THE MATRIX NEVER HAS YOU»
ХИЛАРИ ПАТНЭМ И ТРАДИЦИОННАЯ ФИЛОСОФИЯ

Каковы отношения между нашим восприятием и действительностью? Каковы отношения между сознанием и телом? Каким образом мы придаем значение выражениям нашего языка? На основании чего мы заключаем о мире? Наконец, в чем состоит истинность, и является ли истина свойством или отношением?

Эти традиционные вопросы всегда были и остаются темой философских дебатов. Сразу же бросающаяся в глаза — и в самом деле поразительная — особенность книги Хилари Патнэма «Разум, истина и история» состоит в том, что он показывает *единство* этих вопросов. Усмотрение такого единства носит отчетливо прагматистский характер, поскольку в качестве связующего аспекта выбрана тема рациональности, понимаемой деятельностно, как снятие противоречия между спекулятивными крайностями. Подобный прагматизм характерен для Патнэма на всех этапах его карьеры, и даже регулярно декларируемая им смена позиций означает не перемену направления на противоположное, но скорее радикализацию *Aufhebung* на основе соотношения с практикой. Это относится и к двум блокам исследований, концептуально предваряющим знаменитый аргумент о «мозгах в бочке» — по функционализму и по теории референции, — и к исследованиям по проблеме реализма, наследующим ему.

Философская позиция, которую Патнэм сформулировал и назвал функционализмом в 1960-ых годах, стала очень влиятельной в философии сознания. Функционализм предлагает интерпретацию познавательного процесса как процесса обработки информации — формальной манипуляции репрезентациями или символами согласно некоторым правилам. Согласно этой теории, сознание аналогично вычислительной машине, и поэтому для изучения познания релевантны организация и структура, а не материальный субстрат сознания. Аргумент Патнэма здесь

основан на интуиции о том, что одно и то же ментальное состояние может быть реализовано на различных материальных носителях. Его позднейшие ревизии функционализма связаны с расширением требования о множественности реализаций ментального — в духе свойственного прагматистам плюрализма. Это расширение основано на инверсии исходной интуиции: если различные состояния вычислительной машины или машины Тьюринга могут реализовывать одно и то же ментальное состояние, то сознание не может быть идентифицировано ни с какой конкретной вычислительной машиной. Поддержка Патнэмом сильной формы множественной реализуемости ментального приводит его к пересмотру самого *статуса* психофизической проблемы как ключевой для философии сознания.

Это определяет тот способ, которым Патнэм связывает свою философию сознания со своей философией языка. В этом отношении он следует другому крупному прагматисту XX века — Куайну, давшему мощное орудие для встраивания проблемы языкового значения в очень широкий эпистемологический контекст, рассмотрев в качестве значения референцию — указание на предмет. Патнэм становится одним из родоначальников так называемых «расширенных» теорий референции, согласно которым референциальная связь между родовым термином и его референтом может быть установлена даже в том случае, если говорящий не информирован или плохо информирован о природе референта. Это возможно благодаря тому, что выбор индивида при установлении референции к нему зависит от «каузальной истории» использования данного имени, а она восходит к некоторому акту первоначального именованя. В своей типичной форме этот первоначальный акт устанавливает референцию посредством указательного местоимения или прямого указания на объект реального мира. Поэтому интенционал собственного имени не фиксирован на каком-либо одном представлении в головах говорящих, общем для всех, кто использует это имя как единицу данного языка. Вообще можно сказать, что для каждого термина, введение которого в язык хотя бы частично было основано на остенсивном акте фиксации референции, интенционал этого термина хотя бы частично устанавливается на основе свойств объекта реального мира, первоначально вовлеченных в акт введения термина, хотя выделение объекта про-

исходит не через указание совокупности индивидуальных свойств объекта, а через описание тех обстоятельств, в которых однозначная связь термина и объекта не подвергается сомнению. Поэтому незнание говорящим экстенционала термина или «каузальной истории» термина ведет к незнанию интенционала, но не меняет референциальной ситуации. Согласно Патнэму, границы между экстенционалами терминов устанавливаются более или менее произвольно: в этом смысле содержание существующего определения любого данного термина может быть конвенционально.

Возражения Патнэма против традиционных теорий значения основываются на фундаментальной несовместимости, с его точки зрения, двух важных для них тезисов. Один из них заключается в том, что экстенционал определяет значения общего термина, т. е. то, относительно чего термин истинен. В соответствии с этим тезисом, два термина с одинаковым значением имеют одинаковый экстенционал. Согласно другому основополагающему для таких теорий тезису, существуют определенные психологические состояния людей, соответствующие определенному уровню знания значения.

Таким образом, подход Патнэма противостоит фрегеанской теории значения (и подобным ей семантикам, содержащим менталистские компоненты) в следующем отношении: согласно этому подходу, экстенционал не определен ментальным состоянием говорящего. Фрегеанская теория значения предполагает, что

- (1) знание значения состоит из нахождения в соответствующем ментальном состоянии, и что
- (2) значение определяет экстенционал.

Иными словами, для менталистов знание значения определяет экстенционал термина; ментальное состояние определяет интенционал, а интенционал определяет экстенционал. Патнэм полагает, что такая картина существенно ущербна. По его мнению, само существование реальных сущностей и сама природа референции устанавливают, что значения находятся «не в голове». Между предметами-референтами и их ментальными (или любыми иными) репрезентациями нет *необходимых* связей.

По сути, это очередная переформулировка центрального аргумента аналитической философии сознания и языка середины

XX века, приписываемого позднему Витгенштейну. «Если он справедлив, — пишет Патнэм, — то попытка понять мышление с помощью того, что называют «феноменологическим» исследованием, исходит из неверных оснований, поскольку феноменологи не видят, что описывают внутреннее *выражение* мысли, тогда как *понимание* этого выражения — чье-либо понимание собственных мыслей — представляет собой не *явление или событие*, но *способность*». (Аналогичным образом известный историк феноменологии Г. Шпигельберг задает вопрос: как феноменология все еще возможна, когда она больше не может отрешиться от «ангажированности» в феномены?) Однако классический аналитический аргумент переформулирован с явным постмодернистским акцентом («значения не находятся в голове» — «значения ускользают»). Это обстоятельство — как и все, связанное с постмодернизмом — в равной степени способствовало как его успеху у философствующих гуманитариев, так и ограниченному доверию со стороны профессиональных философов; последние явно предпочли более глубокий вариант выражения сходных интуитивных представлений, который дал Кит Доннелан, и более прозрачный, данный Солом Крипке.

Таков наиболее общий замысел семантики Патнэма (я опускаю здесь эволюцию деталей), непосредственно связанный с аргументом о «мозгах в бочке» — вызовом, брошенным Патнэмом метафизическому реализму.

Метафизический реализм, не допускающий никаких эпистемических ограничений на наши понятия действительности или истины, сталкивается с проблемой объективности или трансцендентности этих понятий — с проблемой возможного отличия того, чем действительность является, от того, чем мы ее считаем, как мы ее воспринимаем или представляем себе. Эта скептическая возможность неоднократно в истории философии была драматизирована различными мысленными экспериментами, классический пример которых — Декартов *malin génie*¹. Патнэм предлагает вариант такого эксперимента, сегодня ставший широко известным по лженаучно-популярному фильму «Матрица»: допустим, что все мы — мозги в бочке, присоединенные к мощному компьютеру, который моделирует наш опыт о

1. Злой гений (фр.). — *Прим. авт.*

мире. Патнэм проводит косвенную атаку на метафизический реализм, выдвигая требование, что это *невозможно*: мы не можем быть мозгами в бочке. Это требование носит отчетливо анти-скептический характер: впрямую оно направлено против скептического аргумента об отсутствии (или невозможности наличия) надежных оснований нашего знания о мире — например, аргумента Чжуан-цзы: философ спит и ему снится, что он бабочка, которая спит и ей снится, что она философ, который спит и т. д. Патнэм пытается разрубить этот гордиев узел дамокловым мечом, заявляя, что невозможность для нас быть мозгами в бочке *априорна*, а скептический аргумент тем самым некорректен.

Аргумент Патнэма связан с его теорией референции. Итак, если мы являемся мозгами в бочке, то мы не можем действительно указывать на внешние объекты, поскольку мы не связаны каузально с реальными вещами в мире; в лучшем случае мы каузально связаны с программированием компьютера. Из этого следует, что, когда мозг в бочке говорит «Это — стол», он делает ложное утверждение постольку, поскольку он намеревается указать на реальные столы, с которыми он каузально связан. То же самое относится ко всем его референциальным утверждениям. Теперь рассмотрим утверждение «Я — мозг в бочке», сделанное мозгом в бочке. Подразумевается, что это референциальное утверждение о реальных мозгах и реальных бочках. Однако нашему мозгу в бочке никогда не были *даны* никакой реальный мозг и никакая реальная бочка; все, чем он располагает — это его машинно-генерированные галлюцинации этих вещей. Ему дано только электрическое возбуждение, но не дано никакого опыта реального мира, в котором эти референты действительны. Но если такие референциальные утверждения, сделанные мозгом в бочке, ложны, из этого следует, что и утверждение «Я — мозг в бочке» также является ложным. Патнэм называет такой аргумент «трансцендентальным» в кантианском смысле, поскольку если мы видим, каковы условия референции, то отсюда *априорно* будет следовать, что мы не можем быть мозгами в бочке. И этого достаточно, чтобы опровергнуть метафизическое реалистическое утверждение, что не может быть никакого *априорного* доказательства против скептицизма.

Философское сообщество восприняло этот аргумент весьма

неоднозначно; его критика может исходить, в частности, из того, что автореферентные утверждения — это в любом случае достаточно специальный класс утверждений, и их характеристики вовсе не обязательно должны распространяться на все референциальные утверждения языка. Кроме того, даже если не рассматривать этот аргумент как некоторую семантическую уловку, его позитивная суть все равно остается не до конца проясненной. Отвечая на эту критику, Патнэм отметил, что реальное значение аргумента состоит в том, чтобы проиллюстрировать «отчужденный» характер метафизического реалистического способа рассмотрения мира. «Отчуждение» здесь состоит в следующем. Такие глобальные возможности, как оказаться мозгами в бочке или жертвами обмана коварным демоном, традиционно рассматривались философами как ставящие глубокие эпистемологические и метафизические проблемы. Но допущение подобных возможностей опирается на идею о том, что мы можем выйти за пределы самих себя, то есть видеть себя и свое состояние с некоторой объективной «точки зрения Бога»; что мы способны, по выражению Томаса Нагеля, на «взгляд ниоткуда». Но, по мнению Патнэма, мы — не Бог (этот популистский аргумент явно апеллирует к тому, что большинство из нас с ним согласится, хотя и по разным причинам), и любая точка зрения — это взгляд *откуда-нибудь*. Поэтому даже если аргумент о мозгах в бочке не представляет собой полное доказательство чего-либо, он тем не менее ясно иллюстрирует возможную смену угла зрения в обсуждении проблемы реализма.

Патнэм характеризует метафизический реализм как позицию, согласно которой существует одно и только одно истинное и полное описание мира, где мир — это полная и фиксированная совокупность независимых от сознания объектов, а истина — это корреспонденция (соответствие) между утверждениями и вещами (с «точки зрения Бога»). Что же создает связь между утверждениями и вещами? Согласно многим современным реалистам, причинные связи — это такие связи, когда мои мысли о столах вызваны реальными столами и т. п., то есть основанные на адекватной корреспонденции. Возражение Патнэма заключается в том, что это далеко не всегда так — например, мои мысли об электронах вызваны учебниками, но они не являются мыслями об учебниках. Реалист ответит, что мы должны

иметь *правильный* вид каузальной цепи. И вот здесь Патнэм с помощью своей теории референции ловит реалистов на циркулярности: мы не будем знать, что данный вид каузальной цепи правилен до тех пор, пока мы не знаем, на что указывают наши слова — а именно это мы и намереваемся установить.

Альтернативная точка зрения, ассоциирующаяся в разных аспектах с антиреализмом, верификационизмом, когерентизмом, прагматизмом, исходит из того, что вопрос «Какие объекты имеются в наличии?» имеет смысл только в пределах соответствующей концептуальной схемы. Здесь, вообще говоря, еще нет ничего неожиданного по сравнению с непосредственными предшественниками Патнэма. Однако из требования, согласно которому наша точка зрения всегда интегрирована в некоторую концептуальную схему, Патнэм делает вывод о том, что истина — это «идеализированная рациональная приемлемость». Основную авторскую нагрузку здесь получает слово *идеализированная*: истинным является не просто то, что было бы рационально принять, исходя из нашей нынешней концептуальной схемы, но то, что мы примем «при эпистемически идеальных условиях» (например, 3000 лет назад теория о том, что Земля плоская, была рационально приемлема, но не истинна). Такая позиция в целом остается в рамках характерного для прагматизма фаллибилизма — восходящего к Пирсу взгляда, согласно которому не существует окончательных верификаторов; нет никаких утверждений, которые не могли бы быть пересмотрены в свете нового опыта. Однако Патнэм предлагает некоторые точки опоры внутри этой бесконечной цепи обоснования нашего знания, обозначает некоторый надежный горизонт — и это горизонт рациональной приемлемости.

Такая трактовка призвана прояснить само понятие рациональности. Важнейшее требование Патнэма состоит в том, что обе наиболее влиятельных концепции рациональности в философии двадцатого столетия неудовлетворительны; фактически, обе опровергают сами себя.

Первая из них — это концепция логического позитивизма, согласно которой утверждения имеют значение лишь в том случае, если они поддаются проверке, и существует установленная процедура для подтверждения (проверки) всех значимых предложений. Эта процедура (множество процедур) — чистый логи-

ческий анализ для утверждений, истинных в силу содержащихся в них символов, плюс эмпирическая проверка в терминах (референциального) содержания для утверждений, заключающих о мире вокруг нас. Рациональность — это не что иное, как соответствие этой установленной процедуре; если мы утверждаем вещи, которые не являются ни аналитическими утверждениями, истинными в силу правил для символов, ни эмпирическими утверждениями, истинными в силу своего референциального содержания, то мы иррациональны. Самоопровержение этой позиции, по мысли Патнэма, состоит в том, что защита самой процедуры не может считаться рациональной: процедура не является истинной в силу содержащихся в ней символов, и при этом она и не истинна в силу некоторого референциального содержания.

Реакцией против позитивизма Патнэм считает релятивизм (куда он относит как Куна и Фейерабенда, так и Фуко и других континентальных философов). Согласно этой позиции, нет никакой установленной формулы или установленной процедуры; что считать «рациональным», может быть установлено каждой культурой (или субкультурой) различным способом. Релятивисты исходят из того, что культурные факторы имеют огромное значение даже в науке и что научные теории зачастую принимаются или отклоняются по причинам политики и финансирования, а не только по причинам чистой проверки в духе позитивизма. Поэтому заключение релятивизма состоит в том, что различные культуры имеют различные стандарты или «парадигмы» рациональности; мы можем оценивать то или иное требование только относительно данной культуры (так, в некоторых культурах рациональны некоторые требования об электронах, а в других культурах рациональны некоторые требования о привидениях, и так далее), и мы не всегда можем перевести наши требования из одной культурной схемы в другую — культурные схемы «несоизмеримы». Самоопровержение релятивизма (который он называет «ментальным самоубийством») Патнэм видит в следующем: формула о том, что является истинным во всех культурах и субкультурах, сама по себе является видом универсального и кросскультурного утверждения, которые релятивисты считают нелегитимными.

Итак, Патнэм старается уйти от эпистемологических крайно-

стей как фундаментализма, так и релятивизма. Однако в глазах фундаменталиста он, несомненно, оказывается релятивистом, так как отрицает наличие и самое возможность *предельного* основания знания — скорее он выдвигает требование *достаточно* основания. Для релятивиста же на этом уровне будет не очень внятен предъявляемый упрек: релятивизм может вообще не апеллировать к допущению о наличии объемлющей, единой эпистемологии. Патнэм же не может — в отличие, например, от того же Фуко — не мыслить в присутствии возможности такого допущения. Мы могли бы сказать, что здесь сама попытка снять глобальное эпистемологическое противоречие, выйти из очерченной им плоскости оказывается важнее результата этой попытки с ее не вполне убедительной концепцией «идеализации». Патнэма, очевидно, не устраивает такое положение, и он предпринимает еще более решительный шаг.

До сих пор в анализе рациональности, как и в анализе реализма, Патнэм применяет даже не постаналитическую, а скорее ортодоксальную аналитическую технику, апеллируя практически исключительно к системам аргументации. В самом деле, до сих пор его протест против абсолютистской концепции трансцендентной истинности или окончательного, предельного основания для наших убеждений до некоторой степени напоминает тезис Куайна об онтологической относительности или критику Дональдом Дэвидсоном различения концептуальной схемы и ее содержания. Но, в отличие от Куайна или Дэвидсона, Патнэм расширяет этот подход на моральную и политическую сферу, доказывая, что крах абсолютистской перспективы в эпистемологии и метафизике делает возможным подтвердить истинностную релевантность морального, политического и других видов управляемого нормами дискурса.

Ход Патнэма состоит здесь в том, чтобы привести доводы против принятой дихотомии между фактическими утверждениями и ценностными суждениями и усмотреть такую позицию, откуда это различие снимается — сохранив, таким образом, единство корпуса человеческого знания. Фактически это возврат на новом уровне к той хорошо известной в истории философии традиции — от Платона до Гегеля, — которая проводила отождествление онтологической и аксиологической проблематики и изначально наделяла бытие ценностным измерением.

Разрыв с этой традицией происходит относительно недавно, когда неокантианцы фрайбургской школы разрабатывают учение о ценностно-нормативном компоненте чистого сознания, выявляемом трансцендентальной рефлексией, и уже у Хайдеггера онтологический анализ оборачивается отвержением аксиологической проблематики. Аргумент здесь примерно таков: если допустить, что бытие (и производные от него образования) само по себе ценностно нейтрально, то тогда требует решения вопрос, каким образом возникает видение вещей с точки зрения их аксиологической значимости, которое пронизывает культуру и которое невозможно отрицать. Отсюда берет свое начало линия поиска истоков ценностного сознания в самом человеке и его культуросозидающей деятельности, чрезвычайно характерная для европейской философии XX века.

Позиция Патнэма, насколько я понимаю, направлена не на отрицание, но на переосмысление этой тенденции. Его возврат к традиции на новом уровне проводится по иному основанию, нежели в светлые времена схоластики, поскольку исходит из критики реализма, и исходным для него оказывается тезис о том, что «каждый факт нагружен ценностью, и каждая из наших ценностей нагружает некоторый факт». *Факт* (или истина) и *рациональность* оказываются взаимозависимыми понятиями.

Патнэм стремится доказать, что даже обычные фактические утверждения (например, «вода состоит из водорода и кислорода») вовлекают некоторые ценностные обязательства, и что оценочные суждения имеют некоторых из особенностей, обычно связываемых с фактическими утверждениями (в частности, оценочные суждения могут быть объективны). Если мы отвергаем метафизический реализм, то неверна картина, согласно которой мы «считываем» факты с мира и измеряем, насколько хорошо наши утверждения соответствуют этим независимым от сознания фактам. Природа процесса познания совершенно иная: мы создаем систему понятий и стандартов, которые будут наилучшим образом соответствовать миру по нашим когнитивным способностям, и оцениваем, что является истинным в соответствии с этими понятиями и стандартами. А если понятия истины или объективности не определены в терминах некоторого отношения к независимым от сознания объектам, но скорее связаны с нашими методами оценки и выбора, то нет пре-

пятствий к тому, чтобы рассматривать моральные и политические требования как самостоятельно объективные и истинные.

Для аналитических философов предшествующего периода характерным считалось отношение к этике, политической и социальной философии и т. п. как к расплывчатым и неясным областям, где изобилуют догадки и спекуляции, но нет четких аргументов, с логической необходимостью ведущих к ясно артикулированным результатам, обладающим истинностным значением. Преодолевая (хотя, разумеется, не первым) эту ограниченность, Патнэм не отрицает различия между науками и такими предметами как этика или политика, но указывает, что абсолютное различие между ними по критериям требований истины и объективности не является надежным; *исторически* больше не может рассматриваться как надежное. Сами точные науки также нормативны: заключения о том, какие теории принимать, или даже о том, какие вопросы ставить, делаются учеными под влиянием ценностей и систем убеждений, и принципиальным образом не отличается от того, что это происходит в моральном и политическом дискурсе. Впоследствии эта позиция привела Патнэма к теории «идеальной коммуникации» Юргена Хабермаса, акцентирующей общественный характер ценностей, как к способу показать, каким образом моральная объективность могла бы быть достигнута в современном обществе, которое является, по общему признанию, субъективистским по характеру; как к способу *согласовать между собой философию языка, эпистемологию и социальную теорию* — задача, вряд ли решенная Хабермасом или Патнэмом, но, безусловно, представляющая собой передний край современной философии, и их заслуги в ее постановке трудно переоценить.

Итак, «сверхзадача» Патнэма здесь состояла в том, чтобы разрушить ту власть, которую естествознание обрело над философской мыслью в XX столетии («появление в культуре философской тенденции, которая была загипнотизирована успехом науки до такой степени, что не могла представить возможности знания и разума вне того, что нам нравится называть наукой — это такой феномен, которого следовало ожидать, учитывая чрезвычайно высокий престиж науки в культуре в целом и учитывая снижающийся престиж религии, абсолютистской этики и трансцендентальной метафизики»). Хотя Патнэм не враждебен науке,

он отклоняет отождествление рационального мышления с научным мышлением и отклоняет идею о том, что наука дает *единственное* истинное описание действительности. Подчеркивая уникальную, ни к чему не редуцируемую ценность самостоятельного философского усмотрения, он, несомненно, остается традиционным философом в лучшем смысле обоих этих слов.