

В.В. Васильев

ТРУДНАЯ ПРОБЛЕМА СОЗНАНИЯ



Содержание

Предисловие
5

Введение
7

Глава 1. На пути к «трудной проблеме»
18

Глава 2. Сёрл: проще простого!
54

Глава 3. Деннет: а вы знаете, что мы зомби?
106

Глава 4. Чалмерс: все решения плохи
152

Глава 5. Назад в будущее
190

Послесловие
248

Литература

253

ПРЕДИСЛОВИЕ

Впервые о загадке сознания я задумался лет в четырнадцать, когда мой отец, Валерий Павлович Васильев, с которым всегда на редкость интересно общаться, удивил меня мыслью, что, по большому счету, человеческий мозг не представляет ничего особенно ценного — просто полтора килограмма органической массы. Я понимал, что это не совсем так, но не придумал, что сказать. И не исключено, что именно тогда я почувствовал интерес к философии. Но философия не терпит любителей, а я не стал бы профессиональным философом, если бы одним зимним утром 1986 г. девочка из параллельного класса не рассказала мне о замечательном философском факультете, который есть в МГУ и на который, по ее мнению, мне обязательно надо поступить — что я и сделал летом того же года.

И разумеется, я хочу поблагодарить своих учителей, а теперь — в течение уже почти пятнадцати лет — и коллег с нашего факультета. Без их понимания и дружбы я никогда не справился бы со своими задачами — ни в этой книге, ни в тех, что предшествовали ей.

С теплотой вспоминаю беседы с А.Ф. Грязновым, который привил мне вкус к аналитической философии. Исключительно ценным было общение с М.А. Гарнцевым, А.Л. Доброхотовым, А.Ф. Зотовым, А.П. Козыревым, В.Ф. Коровиным, А.А. Костиковой, В.Н. Кржевым, А.Н. Красниковым, А.А. Кротовым, В.Г. Кузнецовым, В.Н. Кузнецовым, В.В. Куртовым, Г.Г. Майоровым, В.И. Маркиным, М.А. Маслиным, В.Я. Савреем, К.Х. Момджяном, Е.Н. Моцелковым, В.В. Соколовым, Г.Я. Стрельцовой, Г.Г. Судьным, А.М. Шишковым, О.Д. Волкоговой, А.П. Алексеевым, В.И. Бакиной, Д.В. Бугаев, Е.А. Войниканис, Ф.И. Гиренком, Д.К. Масловым, С.А. Мельниковым, Ю.Р. Селивановым, В.Ф. Титовым, Е.В. Фалевым и многими другими коллегами с философского факультета. Отдельное спасибо нашему декану В.В. Миронову за всестороннюю поддержку.

Я также благодарен коллегам из других институтов: А. А. Гусейнову, Н.В. Мотрошиловой, Т.И. Ойзерману, В.А. Лекторскому, Т.Б. Длугач, Н.С. Юлиной, В.И. Молчанову, А.Н. Круг-

лову, Д.И. Дубровскому, С.А. Чернову, Е.В. Мареевой, Н.Н. Трубниковой и Н. Дмитриевой.

Особо я признателен Ноаму Хомскому и Дэвиду Чалмерсу, общение с которыми позволило мне сформулировать главные идеи этой книги. Очень полезными были и контакты с Д.Деннетом, Н. Блоком, Д.Армстронгом, С. Блэкмор, К. Макгинном, М. Маккинси, Р. Суинберном, Дж. Райдером и Р. Хауэлом.

Большую помощь оказали мне и люди, с которыми я обсуждал отдельные положения данной работы. Прежде всего это Д. Волков, а также Д. Иванов, Е.Косилова, Т. Постникова, Н.Воронов, Б.Ким, Н. Гарнцева, С.Алексеева, Н. Трушкина. И я благодарен — за интересные дискуссии — студентам МГУ, которым я читал спецкурсы по философии сознания.

Я также признателен В.А. Жучкову, А. Мацкевичу, В.К. Шохину и Д.Э. Гаспарян. Они очень воодушевляли меня. И спасибо Насте Данилочкиной, моей маме В.Я. Васильевой, Марине и дочерям — без них я не ступил бы и шагу.

Что касается самой книги, то я решил не дробить ее главы на параграфы, как это часто делается в философских текстах. Я согласен с тем, что это облегчает чтение, и сам не люблю большие разделы. С другой стороны, плохую книгу не спасет даже миллион параграфов, а хорошую прочитают в любом случае. Но главная причина в том, что я пишу в специфической манере. Весь текст этой книги, по сути, целен, и одно утверждение вытекает из другого, так что ее можно рассматривать как одну законченную мысль. При таком положении дел параграфы были бы чем-то искусственным.

Наконец, я должен извиниться перед читателями за возможные и даже вероятные ошибки, которые могли закрасться в мои рассуждения, — хотя я старался, чтобы этого не произошло. В любом случае я надеюсь, что чтение этого текста может принести пользу, развлечь или побудить к каким-то собственным размышлениям. И по крайней мере за некоторые примечания я совершенно спокоен.

21 мая 2008 г.

ВВЕДЕНИЕ

Философия — занимательнейшая вещь. Все мы — я имею в виду коллег по цеху — очарованы ею. Но можем ли мы преподнести публике какие-то «позитивные результаты» своих исследований или должны примириться с мыслью, что философия — это всего лишь замечательная школа интеллекта, помогающая, по крайней мере в идеале, видеть вокруг реальность, а не химеры и замечать в этой реальности то, что достойно быть замеченным? Может ли теоретическая философия как особая дисциплина только ставить проблемы или иногда ей все же удается и решать их? В этой книге я попробую показать, что в некоторых случаях ей это действительно удается. Как ни странно, многие профессиональные философы придерживаются другого мнения. Так что я постараюсь переубедить их.

Эвристичность философии я буду защищать, демонстрируя ее возможности в решении так называемой «трудной проблемы сознания». Но, конечно, эта проблема интересует меня и сама по себе. И разумеется, не только меня — и не только философов, — но и вообще всех, кто в состоянии удивляться (сложно представить себе что-то более загадочное, чем сознание) и размышлять. Впрочем, согласно традиции, к которой я в данном случае присоединяюсь, философ — это именно тот, кто расположен удивляться, а потом размышлять над тем, что его поразило, чтобы понять, как все это устроено. Размышление, нацеленное на понимание, отличается от другой хорошо всем знакомой реакции на удивительное — отмысливания. Отмыслить удивительное — значит без лишних раздумий принять какое-то из имеющихся наготове в культурном пространстве объяснений — и заняться какими-нибудь более насущными проблемами.

Но если мы все же не спешим отмыслить удивительное, а хотим осознанно выбрать то или иное объяснение — или придумать новое, — то чем же может быть обусловлен наш выбор? Можно выбирать «легкие» объяснения, которые соответствуют эмоциональному складу выбирающего. Так часто и делается. Но эмоциональный склад — вещь случайная. У одних один склад, у других — другой. Объяснения, выбранные по таким основаниям, заведомо лишены претензий на общезначимость, а значит, и объективность. А объективность объяснения синонимична его истинности. Итак, если мы хотим подлинных объяснений, мы должны выбирать из

них не те, которые нам подходят, а те, которые выдерживают проверку на истинность. Как проверять на истинность? Анализировать, делать рациональные выводы, сопоставлять эти выводы с фактами, взвешивать за и против. Все эти процедуры обычно называют аргументацией.

Аргументация в таком широком смысле — ремесло не только философов, но и ученых. Такое родство не должно пугать философов. Оно означает, что теоретическая философия — тоже наука: кто думает иначе, тот увлечен чем-то совершенно другим¹. Хотя здесь есть, конечно, немалая проблема. Философия — наука, но иная, отличная от эмпирических или экспериментальных аналогов.

Специфику философии можно почувствовать, если задуматься о различиях философских и нефилософских проблем. Эти различия интуитивно схватываются уже на обыденном уровне, но их концептуализация может потребовать некоторых усилий. К примеру, мы слышим три ювенильных вопроса: что такое молот? что такое вода? что такое вещь? Трудно спорить, что первый вопрос не имеет отношения ни к экспериментальной науке, ни к философии, ответ на второй дает ученый, на третий — философ. Но чем же отличаются эти вопросы? Первый задан о старинном артефакте, определение которого должно даваться в терминах функций, дополненных остенсивными пояснениями², т. е. для ответа на вопрос «что такое молот?» надо указать на какой-то молот и сказать: «Молот — это нечто вроде такой штуковины, предназначенной для обработки металла». Вопрос этот и ответ на него не имеют отношения к экспериментальной науке: она занимается установлением причин, наличие которых в опыте приводит к появлению тех или иных вещей, а функции вещей не могут рассматриваться как достаточные условия их возникновения.

Вопрос о том, что такое вода, имеет несколько иной оттенок. Ответ на него предполагает экспериментальное исследование состава жидкости, которая в повседневной речи именуется водой. Это исследование трудоемко, и неудивительно, что формула воды была установлена лишь в конце XVIII века (не без удовольствия вспоминаю, как в одном немецком учебнике того времени я наткнулся на фразу о том, что, согласно новейшим гипотезам, вода состоит из двух частей водорода и одной части кислорода, хотя, — продолжал автор, — это смелое предположение Лавуазье еще предстоит проверить)³.

Теперь вопрос о вещи. Можно ли представить, что подобно тому, как была установлена химическая структура воды, можно будет когда-нибудь распознать химическую структуру вещи? Или что понятие вещи можно определить через указание на ее функцию? Обе ситуации кажутся просто нелепыми.

Через функцию можно определять артефакты, а вещь — не обязательно артефакт. Что же касается химической структуры, то совершенно очевидно, что ее нельзя установить из-за того, что понятие вещи настолько общее, что говорить о ее конкретном устройстве невозможно.

О чем же тогда можно говорить, отвечая на вопрос, что такое вещь? Ведь вопрос этот не выглядит бессмысленным. Первое, что приходит на ум: можно говорить о правилах употребления слова «вещь» — «вещью мы называем то-то и то-то». Это хороший ответ. Но в нем кроется большая проблема. Ведь слова — такие же артефакты, как инструменты или глянцевые журналы, и если вопрос «что такое вещь?» сводится к вопросу «какова функция слова "вещь"», то этот вопрос ничем принципиально не отличается от вопроса «что такое молот?», который мы только что изгнали из науки и философии.

Читатели должно быть уже догадались, что если сказанное в предыдущем абзаце верно, то философия как научное предприятие — в том понимании, которое я отстаиваю — повисает в пустоте. Разумеется, никто не запрещает уточнять, в каких ситуациях употребляются те или иные слова, и если войти во вкус, это может стать по-настоящему увлекательным занятием, но если это все, то надо помнить, что философия начинается уже с молота⁴. Сомневаюсь, правда, что Джон Остин или тем более поздний Витгенштейн, которые пропагандировали именно такое понимание философии⁵, стали бы отрицать, что переход философии на «молоты» и «наковальни» являл бы собой выход за пределы языковой игры, в которой задействовано слово «философия». А поскольку такой выход запрещен их же философией, так как он приводит к бессмыслице (во многих случаях это и впрямь так), то отсюда следует, что ответы на философские вопросы не являются однопорядковыми с ответами на вопросы о сущности таких артефактов, как наковальни или молоты⁶. Но однопорядковыми они не будут лишь в том случае, если неоднопорядковы сами слова, правила употребления которых мы устанавливаем. Иными словами, есть нефилософские слова и философские слова. Философия выявляет правила употребления философских слов. Это хорошо, вот бы только понять, какие слова являются философскими?

Путеводная нить «вещи» заставляет предположить, что дело в общности. Философские слова — это слова с очень широким употреблением — такова будет наша новая гипотеза. Кому-то она может показаться правдоподобной, но, на мой взгляд, она может представляться истинной лишь в силу привходящих обстоятельств. Непонимание этого момента породило множество путаницы и псевдопроблем. Скажем, слово

«сущее» — наверное, самое общее, если смотреть на его применение. И многие считали, что в философии есть такая проблема, проблема сущего. Думаю, нет такой проблемы⁷.

Почему нет? Да потому, что в слове «сущее» нет одной важной детали, которая присутствует в слове «вещь», да и во множестве других философских слов вроде «материи» или «причины». Чтобы прочувствовать эту деталь, предлагаю немного поразмышлять о некоем возможном мире, тем более, что наши современники Сол Крипке, Дэвид Льюис и др. сделали такие размышления привычными для нас. Возможный мир, в который я предлагаю заглянуть — это мир, изображенный в «Логико-философском трактате» Витгенштейна.

С поздним Витгенштейном мы только что столкнулись, теперь сталкиваемся с ранним. В «Трактате» Витгенштейн рассуждает о том, что язык — это картина мира, и что можно так перестроить обыденный язык, что мы увидим однозначное соответствие между объектами, входящими в состав мира, и знаками, именами нашего языка. Если бы действительный мир соответствовал образу, который внушает читателям Витгенштейн, то значения всех имен — неважно, собственных или общих (последние можно было бы толковать как отсылки к множеству первых)— совпадали бы с объектами, которые они изображают. Мысль, возникающая при произнесении любого имени, должна была бы совершенно уподобляться обозначаемому ей объекту. В противном случае язык не мог бы являться картиной мира.

Так вот, в таком мире могло бы не найтись места для философии, как мы вскоре увидим — поняв, что наличие философских вопросов зависит от одной из особенностей действительного мира, которая не сочетается с нарисованной Витгенштейном картиной (по крайней мере, на первый взгляд — вообще-то с Витгенштейном сочетается все). Мысли, связанные с некоторыми именами в том мире, в котором мы действительно живем, в принципе не могут сводиться к изображению объектов, обозначаемых ими. В этих мыслях есть некий добавочный элемент, нечто примысливаемое нами к тому содержанию, которое мы находим на, так сказать, «объективной стороне», как бы мы ее ни называли — «чувственными данными» или как угодно еще.

Эта привносимая нами добавка в мысли об объектах имеется на самом-то деле при большинстве имен, в том числе при «молоте» и «наковальне». Но там она скрыта за их конкретным содержанием. В более общих именах, таких, как «вещь» или «причина» (но не «сущее», здесь все выхолощено так, что ничего не остается), она выходит на первый план. Поэтому вопросы «что такое молот?» и «что такое вещь?» имеют совер-

шенно разный смысл. Когда я спрашиваю о молоте, я спрашиваю о функции этой вещи (целиком раскрываемой в опыте), когда я спрашиваю о вещи, речь идет о тех моментах, которые привносятся в объективное содержание чувств со стороны Я. Но действительно ли в понятии вещи есть такие привнесения? Похоже, что так, ведь вещь в строгом смысле — это пространственный объект, мыслящийся существующим независимо от нашего восприятия. Но если пространственная сторона вещи доступна в опыте, то ингредиент независимого существования, как нетрудно заметить, и как было замечено еще Юмом, домысливается нами.

Итак, с одной стороны, здравый смысл говорит нам о том, что вопрос «что такое вещь?» — философский. С другой стороны, мы выяснили, что в этом вопросе речь идет о некоторых параметрах, которые мы присоединяем, примысливаем к данным чувственности. Логично поэтому допустить, что философские вопросы как-то связаны с установлением подобных априорных привнесений. Это предположение надо, конечно, еще подтвердить на другом материале. Возьмем вопрос о причинности. Здравый смысл трактует его как философский. Но все тот же Юм четко показал, что, называя какое-то событие причиной другого и полагая это событие в качестве таковой, мы всегда домысливаем к идее этого события нечто, что не содержится в этой идее, а именно представление о том, что данное событие необходимым образом связано с событием, которое оно причиняет.

Вот куда ведут нас философские вопросы. Они заставляют нас присмотреться к субъекту, а не к миру (в том числе и не к «внутреннему миру» наших эмоций и фантазий) — к тому трансцендентальному вкладу, который субъект вносит в явления, налагая некие концептуальные формы на потоки ощущений, составляющих мир опыта.

Я специально выразился по-кантиански, чтобы предупредить читателей, что кантианские ассоциации тут не всегда проходят, Кант утверждал, что наши понятия деятельно формируют мир опыта, здесь же имеется в виду только то, что мы примысливаем нечто к уже имеющимся и каким-то образом упорядоченным данностям. Кант считал, что мы должны до-пускать изначальное упорядочивание нами опыта, так как в противном случае остается загадкой, почему то, что мы априори примысливаем к опыту — например закон причинности — совпадает с тем, что мы реально наблюдаем в опыте, демонстрирующем регулярные связи событий, соответствующие этому закону. Единственной альтернативой своему «коперниканскому» решению он считал теорию предустановленной гармонии⁸ между врожденными законами нашего мыш-

ления и миром опыта - но отвергал ее как совершенно бесперспективный путь, с чем трудно спорить.

Однако в наши дни очевидно, что предустановленная гармония — не единственная альтернатива подходу Канта. Соответствие «врожденных законов мышления», выражающих то, что мы приписываем объектам сверх их непосредственного опытного состава, и гипотетических законов природы — апостериори подтверждающих наши априорные предположения — можно объяснить на основе эволюционной теории, и такое объяснение будет правдоподобнее других⁹.

И тем не менее, ассоциации с кантовской аналитикой, которые могли бы возникнуть у некоторых читателей, конечно, законны. И в этой связи еще несколько уточнений. Во-первых, даже если философия должна заниматься анализом концептуальных привнесений в данности чувственного опыта, философские темы этим, конечно, не исчерпываются. Но именно вопросы о таких привнесениях, как мы увидим, являются фундаментальными, задающими контекст для других философских вопросов — как для служебных и технических, так и для производных и прикладных, а также для мета-вопросов. Во-вторых, философии интересны не все концептуальные привнесения. Некоторые концепты привносятся нами в опыт от случая к случаю, и всегда можно указать их частные эмпирические причины. Философии нечего с ними делать. Другое дело универсальные концептуальные привнесения. Они не зависят от конкретных обстоятельств — не в том смысле, что они повсеместны (хотя некоторые из них могут быть таковыми), а в том смысле, что они являются сущностными компонентами неких типических ситуаций, границы которых, в свою очередь, не произвольны, а очерчены самой природой.

Возьмем, к примеру, такую ситуацию как, восприятие стола, находящегося передо мной. Допустим, у меня возникает мысль, что этот стол неустойчив (что, кстати, — надо же такому случиться — соответствует действительности). Это случайное привнесение, зависящее от наличных качеств стола. Но в этой когнитивной ситуации я могу обнаружить и такие привнесения, которые не связаны с цветом стола, весом стола и т. п., и производятся лишь постольку, поскольку стол — вещь. На это можно возразить, отметив, что качества стола все-таки имеют значение. В самом деле, допустим, что стол, на который мы смотрим, вдруг поменял бы цвет, потом превратился в орхидею, а затем опять принял форму стола, но уже совсем другого по виду. Едва ли мы применили бы в этом случае понятие вещи: мы сказали бы, что это галлюцинация и т. п. Однако справедливость этого возражения не означает,

что в восприятии отсутствуют универсальные (для данной типической ситуации)¹⁰ и априорные компоненты. Ведь функция последних состоит исключительно в том, чтобы доструктурировать имеющиеся в наличии эмпирические данные. Причем структурированность, обнаруживаемая в самих этих данных, рассматривается нами в качестве следствия привнесенного концептуального каркаса. Если же в эмпирических данных отсутствует (или пропадает) такая структурированность, мы просто больше не применяем к ним этот концептуальный каркас. И совершенно необязательно экспериментально устанавливать, какие именно эмпирические данные подходят для того или иного привнесенного нами концептуального каркаса. Ведь это будет понятно из анализа его самого.

Такими-то априорными каркасами и должна заниматься философия. Следуя традиции, их можно назвать категориальными. Значит, философия имеет дело с категориальным анализом. Поскольку категории не исчерпываются конкретными эмпирическими данными, поиск категорий не может сводиться к сопоставлению таких данных. Скорее, эмпирические данные могут быть примерами, рассматривая которые можно выявлять связанные с ними категориальные формы. Выявление это проводится не путем установления корреляций, как в эмпирических науках, а путем анализа, отделяющего опытные содержания от априорных категориальных структур. Сам же этот анализ может проводиться по-разному. Можно, к примеру, отталкиваться от обыденного словоупотребления, исследовать значения слов и смотреть, покрываются ли эти значения эмпирическими ингредиентами. Обнаружение универсальных избыточных аспектов данных значений эквивалентно отысканию той или иной категории.

Нетрудно, впрочем, обойтись и без развернутой лингвистической части — стартуя от типичных когнитивных ситуаций обыденной жизни. Вообразим, к примеру, следующую ситуацию: вы держите на весу некий предмет, скажем, мяч, и собираетесь отпустить его. При этом вы уверены, что мяч упадет и вновь подскочит. Откуда берется это убеждение? Можно ли дедуцировать его из непосредственных данных чувств? Проанализировав эту ситуацию, вы увидите, что ваша убежденность предполагает определенные априорные допущения.

Сам анализ будет иметь следующий вид: вы перечислите компоненты вашего убеждения, затем рассмотрите вопрос о возможности дедуцировать те компоненты, которые непосредственно не даны, из тех, что даны, докажете невозможность этого и сделаете вывод, что уже в исходной ситуации имеется наложение некоей априорной концептуальной фор-

мы или схемы на непосредственные данные опыта. Характер этой схемы выявится по ходу вашего анализа. И сама эта схема задаст границы применения выявленной вами категории, т. е. очертит те самые типичные случаи, для которых она значима.

Хочу подчеркнуть, что категориальный анализ типичных когнитивных ситуаций (необязательно, впрочем, когнитивных — нечто подобное можно производить и при выявлении этических или эстетических категорий) — это лишь начало философского исследования. Самое интересное и по-настоящему сложное начинается после него. Мы должны искать связи категорий, отыскивать их отношения¹¹. А в идеале, возможно, сводить их к некоему первопринципу. Вообще говоря, эта задача отвечает исконному предназначению философии, которая с самых своих истоков понималась как наука о первоначалах.

Впрочем, не будем увлекаться абстракциями. Я хочу пояснить сказанное более наглядными иллюстрациями, а потом, пожалуй, перейду к главному вопросу этой книги — боюсь, что читатели уже заждались этого. Так вот, если учесть, что категории в нашем понимании есть универсальные дополнения чувственных данных, то они могут быть истолкованы как своего рода интерпретации этих данных. К примеру, категория причины мыслится не изолированно от опыта — нет, мы интерпретируем опыт, происходящие в нем события, через нее, т. е. смотрим на события так, что предполагаем, что у них есть какие-то причины. Категория вещи — это интерпретация чувственных ингредиентов в качестве таких, которые сохраняют свое бытие и после прекращения их актуального восприятия. Иначе говоря, категории раскрываются как некие универсальные «онтологические установки». Каждой категории соответствует своя установка. Но вот вопрос: связаны ли они между собой? Как только что было отмечено, выявление подобных связей можно истолковать как одну из важнейших задач научной философии. Но как же выявлять их?

Думаю, механизм здесь все тот же — анализ. Для некоего совершенного разума эти связи — если они имеются — открывались бы в первый же миг его восприятия этих онтологических установок. Но для нас с вами, дорогие читатели, такая возможность исключена. Связи эти приходится кропотливо выискивать. К счастью, они имеют — по крайней мере в ряде случаев — необходимый характер, и их выявление может быть оформлено как строгое доказательство. Ситуация аналогична той, что имеет место в классической геометрии. Высший разум сразу, безо всякого доказательства, постиг бы соотношение катетов и гипотенузы в прямоугольном треуголь-

нике, человеческий же разум постигает его только через доказательство.

Я не случайно поставил философские исследования соотношения онтологических установок сознания (вроде убеждений в существовании причины у каждого события и в независимом от восприятия существовании чувственных данностей¹²) в параллель с геометрическими изысканиями. Научную философию можно было бы даже назвать некоей «геометрией сознания». Впрочем, это словосочетание как бы умаляет нашу область, и я не буду им пользоваться.

Теперь можно наконец презентовать нашу главную тему и указать, каким образом она будет рассматриваться в этой книге. Никто, полагаю, не будет спорить, что мы домысливаем ментальные состояния, т. е. желания, убеждения, образы и т. п. у других людей, а также у иных физических объектов; непосредственно они не даны. Это домысливание имеет универсальный характер для определенного типа ситуаций, и это значит, что перед нами — философская проблема. Мы уже знаем, что границы подобного домысливания должны очерчиваться на основе анализа тех принципов, по которым оно и происходит. Какие-то наметки может, конечно, дать изучение обыденного словоупотребления — к примеру условий буквального употребления слова «чувствовать». Установив их, мы приблизительно поймем условия корректного приписывания ментальных состояний внешним объектам. Но в любом случае нам не обойтись без прямого анализа когнитивных механизмов перенесения собственных ментальных состояний на другие физические сущности, или «одалживания нашего Я», как выразился Кант в одном черновом наброске¹³.

Прояснив эти механизмы, мы сумеем назвать область их применения, т. е. назвать некие физические характеристики, при наличии которых запускаются данные механизмы. Это позволит нам понять характер связи между физическим и ментальным и, соответственно, уточнить онтологический статус сознания. Таким образом, исследование природы ментальных состояний будет обратной стороной исследования законов домысливания таких состояний у других.

На этом пути мы, однако, столкнемся с непредвиденными трудностями. Мы увидим, что онтологические установки, позволяющие приписывать ментальные состояния физическим объектам, противоречат другим установкам. К счастью, эти противоречия окажутся мнимыми и их разрешение позволит вплотную приблизиться к разгадке сознания. Но прежде чем снимать их, нам надо будет представить панораму современных споров о сознании — поскольку все возможные ответы находятся именно в ее границах.

Примечания

¹ Не могу в этой связи не сослаться на Г. Г. Майорова, предлагающего различать три типа философии: «софийную», «эпистемическую» и «технематическую» — см.: Майоров Г.Г. Философия как искание Абсолюта: Опыты теоретические и исторические. М., 2004. С. 34-105. В соответствии с этой классификацией получается, что речь в книге пойдет об эпистемической философии. Впрочем, я не уверен, что это именно разные виды философии. Не исключено, что их можно истолковать, скорее, как ее разные аспекты, в идеале связанные и предполагающие друг друга. В первых строках введения я, правда, противопоставил «позитивную» теоретическую философию и философию как некую технологию взаимодействия с реальностью. Но в конце книги мы увидим, что эта технология может развиваться в том числе на пути решения теоретических вопросов: ключевые проблемы научной философии завязаны на то, чтобы устранить кажущиеся противоречия в установках здравого смысла, а эти установки отражают структуру реальности.

² Потребность в них говорит о том, что в крипкеанской теории референции есть своя правда.

³ Собственно, вот эта книга: Ebert J. J. *Unterweisung in den philo-sophischen und mathematischen Wissenschaften für die oberen Classen der Schulen und Gymnasien*. 4 Aufl. Lpz., 1796. S. 593. Еще в начале XIX в. некоторые — к примеру Гегель — сомневались, что вода — это действительно H₂O.

⁴ Не все, правда, были бы шокированы философией артефактов. Федор Гиренок, к примеру, написал замечательное эссе «Энтелехия кваса», в котором он философствует о том, что материя кваса может вмещаться лишь в форму ковша, но не в бездушные пластиковые стаканчики.

⁵ Одним из любимейших занятий Остина было, как известно, изучение толковых словарей.

⁶ Ответ на вопрос о сущности того или иного артефакта и ответ на вопрос о правилах употребления соответствующего ему слова зачастую совпадают. Отмечу также, что, хотя вопросы о конкретных артефактах нефилософские, вопрос «что такое артефакт?» — философский вопрос и, к примеру, Мартин Хайдеггер мог бы рассказать немало интересного по этому поводу.

⁷ Парадокс, однако, в том, что одна из моих любимых книг — Доброхотов А.Л. Категория бытия в классической западноевропейской философии. М., 1986 — посвящена как раз этой проблеме.

⁸ Kant I. *Kritik der reinen Vernunft*, В 167-168. Кант именовал ее теорией «преформации».

⁹ Разумеется, эволюционные объяснения несвободны от трудностей. Но если говорить о Канте, то проблема в том, что он просто не видел эволюционной альтернативы, не видел алгоритма, механически порождающего целесообразность. См.: Kant I. *Kritik der reinen Vernunft*, В 653-654.

¹⁰ Т. е. для типической ситуации, которую мы называем «восприятие».

¹¹ Ср.: Strawson P.F. *Analysis and Metaphysics: An Introduction to Philosophy*. N. Y., 1992. P. 17-28.

¹² Показать их соотношение можно, к примеру, с помощью такого рассуждения. Задумаемся, почему мы считаем, что окружающие нас какие-то конкретные объекты не исчезают после прекращения их восприятия? Ведь некоторые могут исчезать. Ответ таков: мы считаем, что объект сохраняется в момент, когда мы не воспринимаем его, если, как мы уверены, в его окружении нет причин, которые могли бы его разрушить. А как же быть с возможностью беспричинного уничтожения? Ясно, что мы просто исключаем ее. Подробнее я писал об этом в других местах.

¹³ Kant I. *Gesammelte Schriften*. Bd. 17. В., 1926. S. 679.

Глава 1 НА ПУТИ К «ТРУДНОЙ ПРОБЛЕМЕ»

Мы только что обсудили такое понимание задач философии, согласно которому главной из них является анализ соотношения универсальных онтологических установок. Одной из таких установок является убежденность в существовании «других сознаний». Уже беглого взгляда оказалось достаточно, чтобы понять: прояснение ее природы совпадает с уточнением соотношения ментального и физического, а значит и с выяснением сущности сознания. Это — одна из давних забот философии: ее можно расслышать уже в знаменитом призыве к самопознанию, озвученном при самом ее рождении,

В этой главе нам предстоит понять, не вышла ли из моды эта тема. Некоторые философские проблемы и впрямь устаревают. Например, после Юма и Канта утратили актуальность доказательства бессмертия души и доказательства бытия Бога¹. Если какая-то философская проблема выходит в тираж, это верный знак того, что она была решена, — неважно, позитивно или негативно. И это, кстати, свидетельствует о том, что в философии есть прогресс². Так вот, может быть и вопрос о соотношении сознания и мозга, ментального и физического уже не интересует философов так, как раньше? Это означало бы, что и он нашел свое решение. И наоборот, повышенный интерес к данной теме наводит бы на мысли о каких-то теоретических трудностях, которые не удается преодолеть. Ведь концептуальная путаница — подлинный источник вдохновения для всех нас.

Но прежде чем понять, вдохновляет ли современных философов вопрос о природе сознания, надо установить, а сохранились ли они вообще? Со своих студенческих лет, т. е. с конца 80-х годов, я слышал рассуждения о кризисе философии или вообще о прекращении ее существования. Вспоминаю, к примеру, как наш знаменитый филолог, В.В.Иванов, на одной из лекций (которые, к слову, отличались весьма нетривиальной мозаичностью) объявил, что недавно он был в Германии и в каком-то баре — кто бы сомневался — повстречал последнего философа. Досаду Иванова можно понять, и он не одинок в своих оценках³. Хотя, конечно, не все столь радикальны. Более взвешенными выглядят суждения, говорящие о том, что философы не столько исчезли, сколько изменились, исчезли лишь классические философы. Показательной в этом плане была позиция Ричарда Рорти⁴. Он считал, что эра Фи-

лософов закончилась, и началось время философов — с маленькой буквы. Раньше философы придумывали системы, решали кардинальные вопросы, а теперь им надо приучить себя к скромной роли посредников, налаживающих коммуникацию представителей различных дискурсов. Философия не может формулировать проблемы и разрешать их.

Хотя к мнению Рорти нельзя не прислушаться, не стоит все же забывать, что он был изгоем англо-американской философии, из которой он вышел. И сам этот факт говорит о том, что большинство англо-американских мыслителей придерживается другого понимания задач философии. Это большинство считает, что философия способна на позитивные свершения. Иными словами, это большинство понимает философию в классическом смысле. И это не случайно. Ведь так называемая «аналитическая философия», к традиции которой относит себя подавляющая часть англоязычных мыслителей, и есть наследница классической философии, ориентированной на строгость аргументации, признание у себя специфической проблемной области и допускающей возможность реального прогресса в решении этих проблем⁵. Одним словом, беспокоиться за классическую философию в начале XXI века нет оснований. Хотя в Германии и Великобритании и в самом деле наблюдались некие кризисные явления (британцы должны поблагодарить М. Тэтчер, в свое время урезавшую академические ставки⁶), это означает лишь то, что центр мировой философии переместился в какое-то другое место.

В сдвиге философского центра нет ничего удивительного, и XX век дал немало примеров такого рода эрраций: в 20-е и 30-е годы этот центр был в Вене, потом — в Кембридже Рассела, Мура и Витгенштейна, затем на десятилетие сместился в Оксфорд Райла, Айера (ученика Райла и провозвестника «венской» мудрости), Остина и Стросона. В 60-е годы фокусом философских дискуссий стал Париж Фуко, Делеза и Деррида, а вот затем, промелькнув во Франкфурте, философский центр оказался за океаном. Сначала он был в американском Кембридже Куайна, Хомского и Патнэма, а потом и вовсе исчез, расползаясь по всей Северной Америке — хотя некоторые считают, что главный философский город в начале XXI века все же есть, и это Нью-Йорк⁷. Возможно, и это было бы даже логично — столица мира, где, как говорится, даже воздух соткан из культуры, может быть и столицей философии, но для нас сейчас важнее другое: трудности с определением «географического центра» современной философии связаны не с отсутствием периферии, а с ее интенсивным присутствием, нарушающим четкость метафоры окружности.

В общем, те, кто говорит о закате философии, несколько драматизируют ситуацию. Реальная картина иная. Бум, горячка, невероятный расцвет — вот истинное положение вещей. Так что к реальности ближе австралийский мыслитель Дэвид Армстронг, как-то сравнивший XX век в философии с золотой эрой Платона и Аристотеля.

Причины высокой востребованности философии, ставшей особенно очевидной в конце прошлого столетия, разнообразны. Отчасти они вписываются в уже очерченную ранее картину. Как подметил журналист Томас Фридман, наш мир становится все более «плоским», все более коммуникабельным, и это неизбежно сталкивает различные научные и культурные дискурсы, порождая потребность в людях, которые могут найти точки их соприкосновения и сформулировать рациональные решения ранее невиданных проблем. Эта общая причина особенно действенна в одной из научных областей.

Речь идет как раз об исследованиях сознания. Проблема сознания была одной из центральных тем философии XVII-XIX столетий, но в начале XX века, в связи с кризисом традиционной метафизики и триумфом бихевиоризма, она отошла в глубокую тень. Сознание стали считать феноменом, недоступным объективному научному исследованию. Неважно, какова природа сознания, субъективных ментальных состояний, важно, что научно изучать можно лишь поведение. Причем речь шла даже не об установлении корреляции ментальных состояний с поведением. Нет, субъективные ментальные состояния вообще исключались из рассмотрения. Бихевиористов интересовала другая корреляция: поведения и внешних причин, которые его вызывают, или, иными словами, связь реакций и стимулов.

К середине XX века бихевиористское движение возглавил Б.Ф. Скиннер. Его блестящие эксперименты с крысами и голубями показали, что, умело подкрепляя реакции, можно формировать у животных сложное поведение. Воодушевившись такими результатами, Скиннер распространил бихевиористский подход на людей⁸ и, в частности, на важнейшую разновидность человеческого поведения — речь. Именно этот шаг, как показали дальнейшие события, стал поворотным пунктом в исследованиях сознания в XX веке. Скиннер опубликовал свои размышления о языке и речи в книге 1957 г. «Вербальное поведение»⁹, а в 1959 г. в журнале *Language* появилась рецензия на эту работу молодого лингвиста Ноама Хомского¹⁰.

Хомский нанес сокрушительный удар по теории Скиннера. Скиннер считал, что ребенок научается говорить вследствие влияния на него вербальных стимулов и разного рода

подкреплений. Хомский, однако, показал, что классические бихевиористские понятия «стимул», «реакция» и «подкрепление», вполне применимые к крысам или голубям, которые сидят в коробке и нажимают рычаги, плохо работают, когда мы говорим о вербальном поведении¹¹. Скал-сем, стимулы к вербальному поведению могут быть не внешними, а внутренними. Сила вербальных реакций тоже не может быть измерена по внешним проявлениям. Не обязательно внешним является и подкрепление. Язык, к примеру, можно учить ради самого обучения, так же как решать задачи можно просто ради удовольствия. Тем самым Хомский показал, что если бихевиористские понятия «стимул», «реакция» и «подкрепление» и могут использоваться при объяснении научения языку, то при этом приходится лишать их «публичного», наблюдаемого содержания, на допущении которого основана вся идеология бихевиоризма. Поэтому Хомский вполне резонно заключил, что Скиннер довел бихевиоризм до самоотрицания¹².

Но мало разрушить какую-то доктрину — а критика Хом-ского пришлось настолько ко времени, что именно с ней связывают крах бихевиоризма: Давид победил Голиафа (не исключено, правда, что это преувеличение: бихевиоризм разрушался под собственной тяжестью и многие его критики, такие, как Д. Деннет, вышли из него самого, из его интеллектуальной среды), — надо еще предложить альтернативные решения. Позже Хомский упрекал себя за то, что в упомянутой рецензии он крайне осторожно высказывался на этот счет, хотя альтернатива у него уже была. Прежде чем сказать о ней, задумаемся о том, какой она в принципе может быть. Если ребенок научается речи не только в силу внешних физических влияний, то это означает, что он привносит в этот процесс что-то и от себя. Но что?

Первая мысль: ребенок привносит от себя некие базовые эмоциональные, а также когнитивные принципы вроде индуктивной склонности переносить на новые случаи то, что он знает о похожих случаях, зафиксированных в его прошлом опыте. К примеру, он слышал выражение «Вот мяч!» от родителей, и теперь, увидев похожий предмет, он повторяет эту фразу. Если это не мяч, а воздушный шарик, то родители поправляют его и, сравнивая образы мяча и шарика, он замечает отличия и переносит свои выводы на будущее. Запомнит он свою ошибку и при поправленной синтаксической неточности. Так ребенок постепенно научается правильно выражаться.

Проблема, однако, в том, что бихевиористы в действительности не отрицали наличия внутренних когнитивных механизмов. Но они были уверены, что, будучи механизмами об-

щего действия, они могут быть вынесены за скобки при установлении корреляций стимулов и реакций. Вполне возможно, что в этом и состояла их главная ошибка. Но если бы речь шла только о таких механизмах, выпады Хомского против Скин-нера утратили бы свою остроту. Хомский, однако, противопоставлял бихевиористским объяснениям формирования человеческой речи гораздо более специальные тезисы. Он утверждал, что общих индуктивных механизмов — врожденность которых могут допускать и бихевиористы — совершенно недостаточно для научения языку. Суть в том, что человеческая речь ориентирована не на повторение, как было бы в случае ее индуктивно-эмпирического формирования, а на порождение и понимание неограниченного количества новых предложений. Выходит, наша речь содержит существенный креативный компонент, который должен иметь внутреннее происхождение. Кроме того, в процессе обучения языку ребенок получает лишь очень фрагментарную информацию о его синтаксисе из внешней среды, и никакая индукция, по Хомскому, не помогла бы ему усвоить грамматические правила - а тем не менее дети успешно усваивают их до появления в начальной школе¹³.

Рассуждения Хомского об информационной недостаточности вербальных стимулов оказались особенно влиятельными. И в самом деле, способность детей быстро усваивать язык кажется почти чудом. Но это чудо можно объяснить, признав существование врожденных специализированных механизмов, настраивающих детей на понимание речевых актов и генерацию вербального поведения. Поскольку кажется, что именно грамматика является камнем преткновения для бихевиористских или эмпиристских (как предпочитает называть их Хомский) объяснений научения языку, логично допустить, что люди изначально наделены неким грамматическим арсеналом. Факт многообразия языков подсказывает, что этот врожденный грамматический арсенал должен иметь абстрактный характер и содержать лишь общие грамматические принципы. Его-то Хомский еще в ранний период и назвал «универсальной грамматикой».

Универсальная грамматика, конечно, должна быть реализована в каких-то физических структурах мозга, которые в принципе доступны внешнему наблюдению. Так что кажется, что даже в этих обстоятельствах бихевиоризм, пусть и в расширенном толковании, может устоять. Однако на деле внешнее наблюдение за подобными структурами до такой степени затруднено, что проще задействовать при их анализе внутренний опыт. Таким образом, теория Хомского дала сильнейший импульс к возрождению так называемого мента-

лизма, т. е. внимания к внутренним аспектам психических состояний.

Многие в наши дни согласны, что это была благотворная революция. Другое дело, не зашел ли сам Хомский слишком далеко. В мою задачу сейчас не входит ни отслеживание истории концепции универсальной грамматики (в прошлые эпохи или у самого Хомского, который менял свои взгляды¹⁴), ни детальный анализ его аргументов. Отмечу лишь, что тезис о грамматической неполноценности внешних стимулов кажется мне спорным. Заметим, что никто не пытается всерьез отрицать положение, что ребенок научается значению конкретных слов из опыта¹⁵ (хотя в них могут или даже должны быть интегрированы априорные компоненты, не выходящие, однако, за пределы общих когнитивных принципов). Но, в конце концов, слов гораздо больше, чем правил грамматики. Почему же мы исключаем возможность эмпирического усвоения последних, тем более что речь идет не о грамматиках конкретных языков (которые тоже не могут быть врожденными), а о самых общих принципах? И главное, как доказать, что ее следует исключить? Однажды в письме я спросил об этом самого Хомского. Его ответ свелся к трем пунктам.

Во-первых, он упомянул достоверные случаи, когда дети вообще не получали практически никакой информации от взрослых, но, общаясь друг другом, словно бы изобретали - а в действительности, по Хомскому, извлекали из врожденных структур — язык с развитым синтаксисом¹⁶. Во-вторых, наличие врожденной грамматики доказывается хотя бы тем, что, к примеру, внучка Хомского усвоила язык, а ее птичка, которая слышала то же, что и она — нет. В-третьих, никто, по Хомскому, не смог убедительно показать возможность другого объяснения усвоения языка и показать, как оно было бы возможно на основе общих когнитивных механизмов. Однако, на мой взгляд, представленные им аргументы (сходные доводы он, разумеется, приводил и в других текстах) в принципе несложно оспорить. Птичка могла не усвоить язык, потому что ей недоставало общего интеллекта, равно как и потребности общаться с людьми, дети изобрели такой синтаксис, который позволил им с наименьшей тратой сил описывать мир с характерной для него онтикой и т. п.¹⁷

Но, повторяюсь, сейчас нас интересует не оценка теории Хомского - она, как мы увидим, не имеет отношения к «трудной проблеме сознания», которой посвящена эта книга, — а описание ее влияния, через которое мы выйдем к осмыслению этой проблемы. А влияние теории Хомского было громадным. Одни лингвисты бросились систематизировать универсальные грамматические структуры, другие занялись

детальным изучением разных языков, чтобы за внешними различиями отыскать их глубинные сходства и т. п.¹⁸ Но воздействие идей Хомского не ограничилось лингвистикой. Они революционизировали и науки о сознании, и вот почему. Хомский трактовал способность людей к научению языку как относительно автономный модуль, лишь опосредованно связанный с другими человеческими способностями. Такое понимание оказалось очень перспективным. Ведь оно подталкивало к вопросу: а не состоит ли все человеческое сознание из аналогичных специализированных модулей?

Сам Хомский осторожно отвечал на этот вопрос, но все же отвечал положительно¹⁹. Его ученики, как всегда, оказались более радикальными. В 1983 г. американский философ Джерри Фодор опубликовал книгу «Модульность сознания», которая стала своеобразным манифестом новой теории²⁰. Правда, Фодор отрицает, что сознание состоит исключительно из модулей — в нем есть и немодульная, интегральная часть²¹. Однако и эта позиция показалась слишком уж компромиссной другому последователю Хомского - Стивену Линкеру. В книге «Как работает сознание» (1997) он придал тезису модульности универсальный характер, чем навлек на себя резкую критику Фодора, который вообще славится острым умом, грубоватым юмором и агрессивностью по отношению к коллегам. Нападая на Линкера, Фодор опубликовал монографию «Сознание так не работает» (2000), вынудив своего оппонента на ответный ход: статью «Так как же работает сознание?» (2005)²². Эти споры вокруг идейного капитала Хомского (сам он заявляет, что вопрос о модулях слишком запутан, чтобы вмешиваться в полемику) доказывают, что в нем есть нечто достойное обсуждения. Речь идет вот о чем. Во-первых, модульная теория сознания открывает дорогу для объединения психологии и эволюционной биологии, во-вторых, она оптимально подходит для реализации проектов так называемой когнитивной науки. В-третьих, она имеет прямые выходы на нейронауку.

Если эти строки сейчас читает человек, который только приступает к исследованиям сознания, он может поначалу не оценить значимость каждого из перечисленных выше пунктов. Но, уверяю, каждый из них дорогого стоит. Взять, к примеру, проблему объединения теории эволюции и психологии. Об этом мечтал еще Дарвин, но до самого последнего времени все подобные попытки оставались чисто спекулятивными. Теория эволюции гласит, что изменчивость особей при равных стартовых условиях и конкурентном доступе к жизненно важным ресурсам приводит — когда речь идет о больших популяциях — к выживанию тех, мутации которых в наиболь-

шей степени способствуют доступу к этим ресурсам. Если изменения сохраняются у потомства, то со временем особи все лучше и лучше адаптируются к окружающей среде или же занимают новые ниши.

Рассуждая с этих позиций о человеке, можно с уверенностью утверждать, что если сознание — некая унифицированная инстанция, то все наши частные способности и поведенческие реакции должны быть продуктом опыта и культурных различий, и союз психологии и дарвиновского эволюционизма лишен больших перспектив. Напротив, при детерминации этих способностей и реакций различными генами, т. е. при модульном строении сознания, открывается возможность спрашивать об их адаптивной функции. Скажем, допущение существования относительно автономного лингвистического модуля, или «органа», позволяет ставить вопрос о выгоде, которую могла принести «лингвистическая мутация» человекоподобным особям. И ответ на этот вопрос не выглядит чем-то нереальным. Появление речи дало нашим предкам возможность более эффективно распространять важную информацию, и это способствовало их выживанию.

Аналогичную процедуру «обратной инженерии» можно провести и с другими модулями — отвечающими за эмоциональные реакции людей на родственников и друзей, за идентификацию вещей, за ориентацию в пространстве и многие другие вещи. Эволюционистские объяснения подчас имеют выход на практику: некоторые модули, возникшие в первобытной среде, могут давать сбои в условиях современной цивилизации, и понимание их изначальной функции вполне может указывать пути смягчения возникающих коллизий²³.

Самым известным проponentом эволюционистских объяснений модулей сознания является Линкер. Но этот подход нравится далеко не всем и, в частности, Хомский весьма скептичен относительно его преимуществ. Один из контрдоводов состоит в том, что из теории эволюции не вытекает, что все наследуемые изменения должны быть адаптивными²⁴. Такие изменения, к примеру, могут лишь сопутствовать другим изменениям. И если это верно для модулей, то из существования некоего модуля нельзя заключать, что он приносил какую-то выгоду нашим первобытным предкам. А это означает, что метод обратной инженерии сознания лишен эвристичности.

Этот аргумент может показаться сильным. В самом деле, допустим, что организм претерпевает мутацию, приводящую к появлению у него двух сопутствующих признаков А и В, причем А выгоден для организма, а В вреден. Выгода от А, однако, перевешивает вред от В, и эти признаки сохраняются в результате действия естественного отбора. В итоге получает-

ся, что «отбираться» могут даже вредоносные признаки. Максимум, о чем можно говорить в этой связи, так это о том, что некоторые признаки допускают адаптационное толкование. Тонкость, однако, в том, что модули, точнее те поведенческие эффекты, которые они порождают, по всей видимости, относятся к числу именно таких признаков. Ведь модули кажутся относительно независимыми друг от друга, а значит, маловероятно, что они детерминируются одними и теми же генами. Это, в свою очередь, делает маловероятными мутации, приводящие к одновременному появлению двух или большего количества модулей.

Оставим, однако, этот вопрос в стороне и поговорим о другом следствии модульной теории сознания, относительно полезности которого согласны почти все. Я имею в виду то обстоятельство, что эта теория прекрасно согласуется с основоположениями когнитивной науки. И теперь я должен сказать несколько слов об этих основоположениях. Предельно кратко их можно выразить следующим тезисом: человеческое сознание аналогично компьютерной программе, а мозг аналогичен самому компьютеру, т. е. его аппаратной части.

На первый взгляд эта мысль может показаться странной, но она прочно завладела умами наших современников. Так что стоит присмотреться к ее истокам. Первый современный компьютер — в теоретическом исполнении — был создан английским математиком Аланом Тьюрингом в 30-е годы XX века. Не вдаваясь в подробности, можно сказать, что он разработал модель механического устройства, способного выполнить любой алгоритм, т. е. ту или иную последовательность элементарных шагов, а в итоге решающего задачи, которые могли бы казаться специфичными только для человеческого интеллекта. Одним словом, Тьюринг показал, что самые примитивные машины могут «мыслить». Приведу пример такого мышления. Допустим, наша машина должна отследить определенную последовательность на ленте нулей и единиц, скажем такую, где 0 идет за 1. Эта задача совсем простая, но она соответствует тому, что понимается под мышлением, ведь мышление предполагает сопоставление и анализ данных. Так вот, машина может выполнить такую интеллектуальную задачу. Для этого достаточно расписать (а затем и механически реализовать, что совсем несложно) ее «внутренние состояния», которые будут определять ее действия. Первое из них устанавливает, что появление на ее считывающем устройстве 0 должно оставаться без последствий, тогда как появление 1 должно приводить к переходу ее во второе состояние, в котором появление 1 должно оставаться без последствий, но появление 0 должно останавливать машину и,

к примеру, давать команду на печать сообщения «искаемая последовательность найдена». Нетрудно заметить, что «внутренние состояния» есть связки импликаций типа: «если А то С», а «если В то D и E», где А, В, С, D и E могут быть не только элементарными данными на входе и выходе, вроде 0 или 1, но также «действиями» или новыми «внутренними состояниями». Описания алгоритмов, выполненные в таком виде, можно назвать программами.

Но какое, можно спросить, отношение эта компьютерная схематика имеет к человеческому сознанию? Конечно, мы только что видели возможность проведения некоей аналогии действия Машины Тьюринга с мышлением²⁵, но ее можно истолковать так, что это человеческое мышление может быть машинообразным, а не наоборот²⁶. Однако дело вовсе не только в упомянутой аналогии. Еще в статье «Сознание и машины»²⁷ (1960 г.) знаменитый гарвардский Протей аналитической философии Хилари Патнэм показал, что между сознанием как совокупностью ментальных состояний и программами имеется гораздо более глубокое родство. Суть в том, что ментальные состояния, такие, как желания и убеждения, функционируют не разрозненно, а сплетены в некую сеть, напоминающую программные связки импликаций. К примеру, если я хочу выпить мохито, то появление передо мной бокала с этим коктейлем приведет к тому, что я возьму бокал и выпью мохито, но лишь если я убежден, что этот бокал не галлюцинация, что я имею право взять этот бокал, если я уверен, что там и впрямь находится мохито, и т. д. Иными словами, внутреннее состояние, при котором я протяну руку за бокалом с мохито, является следствием определенной совокупности предпосылок, так же как внутреннее состояние, при котором Машина Тьюринга совершает то или иное действие, определяется целым рядом условий. После того как я выпью мохито, я перейду в новое «внутреннее состояние», для которого будут характерны новые типы поведенческих реакций. Это и правда похоже на Машину Тьюринга.

Патнэм как раз и подчеркнул данное сходство. Он объявил мозг аналогом материальной составляющей Машины Тьюринга, а сознание — аналогом ее программы. И этот факт стал одной из отправных точек когнитивной науки. Пафос когнитивистов был связан с тем, что понимание сознания как программы позволяет абстрагироваться от аппаратной части, т. е. от мозга (ведь одна и та же программа может быть исполнена на разных hardware, причем детали архитектуры этой материальной основы несущественны), и научно исследовать виды ментальных состояний, которые в таком случае приравниваются к вычислительным состояниям Машины Тьюринга.

Со временем, впрочем, выяснилось, что уверенность некоторых когнитивистов в том, что они могут обойтись без нейро-науки, была не вполне обоснованной²⁸. В конце концов, и рассуждения об одной и той же программе, исполняемой на разных компьютерах, это тоже идеализация. Программа, написанная для IBM PC, не пойдет на Mac. Детали архитектуры могут иметь критическое значение.

Обнаружилась и другая неприятная для когнитивизма вещь - абстрактность компьютерной метафоры. Резким критиком когнитивизма в этом направлении стал сам Патнэм, разочаровавшийся в собственных идеях²⁹. Причем с годами эта критика становилась все острее. В поздних статьях он уже прямо говорит о «химере» когнитивной науки³⁰. И в самом деле, хорошо рассуждать о том, что, скажем, желание сходить в Метрополитен-музей — это вычислительное состояние, но попробуйте привести пример реальной программы, вычисляющей место этого состояния в ряду других ментальных состояний! Никто и никогда еще не сделал этого, уверен Патнэм.

Критика Патнэмом своих ранних идей понятна, но стоит обратить внимание, что она основана на холистском, в куай-новском смысле, понимании ментальных состояний: все такие состояния или, по крайней мере, пропозициональные установки, в частности убеждения и желания, зависят от всех других подобных состояний — а также, по Патнэму, от разного рода внешних факторов - - и индивидуализируются лишь в отношении к целому. При таком подходе задача программного описания конкретных ментальных состояний и в самом деле выглядит почти нерешаемой. Модульная теория сознания, к которой мы возвращаемся, рисует, однако, лучшие перспективы³¹.

Чтобы оценить ее преимущества, достаточно обратить внимание на ее основную черту. Модули мыслятся как относительно автономные блоки, каждый из которых может иметь свою программу, нацеленную на решение частных задач, — и не обязательно сложную. Так что здесь вполне реально построение математических моделей, а возможно, и механическое воспроизведение тех или иных блоков. Надо сказать, что в этой области сделано уже немало. Довольно впечатляющими были попытки построить математические модели зрительного восприятия. И тут наверняка будут новые достижения.

Еще один довод в пользу модульной теории сознания состоит в том, что она, по-видимому, подтверждается данными экспериментальной психологии и неврологии. Если взять то же зрительное восприятие, то его модульный, относительно автономный характер демонстрирует хотя бы то, что всем из-

вестные зрительные иллюзии, когда одинаковые цвета кажутся разными, когда равные отрезки кажутся отрезками разной длины и т. п., не исчезают и после того, как мы рационально убедили себя, что это всего лишь иллюзия³². Другие подтверждения модульности дают клинические случаи, связанные с повреждениями тех или иных участков мозга, при которых оказывается, что человек, к примеру, может идентифицировать столы и стулья, но не в силах узнавать людей или сохраняет высокий интеллектуальный коэффициент, но утрачивает способность грамматически правильно выражать свои мысли и т. п.³³ Если бы сознание было устроено холистически, т. е. являло собой единую сеть ментальных установок, то частичные повреждения мозга могли бы легко компенсироваться и подобных драматических феноменов не наблюдалось бы. Их существование показывает, что психические функции жестко специализированы и привязаны к определенным участкам нервной системы, пластичность которой не стоит преувеличивать.

И вот в этом, в указании на тесную связь определенных классов ментальных состояний и конкретных подсистем мозга, состоит еще одно несомненное достоинство модульной теории сознания, хотя я уже говорил, что не так давно некоторые когнитивисты назвали бы это скорее недостатком, чем достоинством. Наука о мозге до последнего времени была настолько неразвита, что было просто странно искать союза с ней. Но сейчас ситуация резко изменилась.

Эти радикальные изменения были связаны с целым рядом обстоятельств, не в последнюю очередь с уточнением биохимии нейронных процессов и появлением новых способов получения детальной картины работы мозга при выполнении им самых разных задач.

Но, разумеется, дело не только в fMRI и других современных методиках³⁴. Не меньшее значение имело то, что во второй половине XX века нейронаука привлекла внимание целой армии незаурядных ученых. Пионером изучения сознания с современных нейробиологических позиций стал Фрэнсис Крик. Разгадав в 50-е годы вместе с Дж. Уотсоном «загадку жизни» — определив структуру ДНК³⁵, — в последующий период он обратился к решению загадки сознания на базе конкретных нейронаучных изысканий. Крик и его коллега К. Кох начали настоящую охоту за «нейронными коррелятами сознания»³⁶.

Отмечу, что их интересовало не сознание в широком смысле — как совокупность ментальных состояний и механизмов (привязка некоторых из них к определенным участкам мозга к тому же давно установлена), — а сознание в узком смысле,

как состояние, противоположное состоянию сна без сновидений или обморока, и даже не сознание в узком смысле как таковое, а тот его аспект, который связан с восприятием зрительной информации. Они пытались понять, какие нейронные процессы отвечают за появление визуальных контентов³⁷.

Поначалу, еще в 1990 г., Крик и Кох, опираясь на экспериментальные данные и гипотезы немецких нейрочеловек В. Зингера, Р. Экхорна и др., предположили³⁸, что, поскольку визуальное сознание интегративно, т. е. предполагает объединение таких составляющих, как цвет, форма, движение и т. п., за обработку которых, однако, отвечают разные группы нейронов, то нейронными коррелятами сознания могут оказаться те процессы в соответствующих «визуальных» участках переднего мозга, которые синхронизированы друг с другом, — а синхронизация, полагали они, происходит через высокочастотные колебания. Между тем в 2004 г. Кох заявил, что «сегодня ни Фрэнсис, ни я больше не считаем, что синхронизированные импульсы являются достаточным условием нейронных коррелятов сознания»³⁹. К отказу от этой гипотезы их привели экспериментальные трудности⁴⁰. Взамен Кох выдвинул тезис, что синхронизация нейронных импульсов играет более скромную роль — одного из факторов, способствующих победе какой-либо из конкурирующих коалиций нейронов. Победитель получает возможность влиять на разного рода подсистемы в мозге, в частности на центры, отвечающие за речь и планирование⁴¹. И в этой связи Кох отсылал⁴² к модели сознания, предложенной американским когнитивным ученым Б. Баарсом⁴³.

Речь идет о так называемой «теории глобального рабочего пространства». Баарс утверждает, что функциональная архитектура мозга напоминает место вокруг доски с информацией⁴⁴ или даже, скорее, театр. Здесь тоже есть некое подобие сцены, попадая на которую в результате деятельности множества невидимых агентов те или иные паттерны или контенты становятся доступными для использования другими подсистемами и модулями. Освещенная сцена есть не что иное, как аналог сознания. Нами сознаются именно те процессы, которые оказываются на этой сцене, попадают в глобальное рабочее пространство или, иными словами, «транслируются в мозге».

Я только что упомянул, что Кох соглашается интерпретировать свою концепцию в терминах глобального рабочего пространства. Но ценность исследований Коха и Крика в том, что они пытались насытить абстракции Баарса конкретным нейробиологическим содержанием. Ведь один из главных вопросов, возникающий в связи с этой теорией, таков: где находит-

ся глобальное рабочее пространство?⁴⁵ Кох предлагает версию ответа. Опираясь на впечатляющие массивы экспериментальных данных, он утверждает, что, если говорить о визуальной информации, ворота в глобальное рабочее пространство — и в сознание — находятся в нижней височной коре мозга⁴⁶.

Разумеется, Крик и Кох не обладали монополией на поиски нейронных коррелятов сознания и нейробиологическую интерпретацию баарсовской модели глобального рабочего пространства. Близкую по духу теорию — как признает сам Кох — развивает, в частности, французский нейрочеловек С. Деан с коллегами⁴⁷. Но Баарс симпатизирует разработкам нобелевского лауреата Дж. Эдельмана⁴⁸.

Согласно Эдельману и Дж. Тонони⁴⁹, мозг является местом «дарвиновского отбора» конкурирующих групп нейронов, в результате которого в нем формируется так называемое «динамическое ядро», которое при желании можно истолковать как субстрат глобального рабочего пространства, т. е. информационной доски в мозге. Это ядро — некий «функциональный кластер» — образуется тогда, когда различные участки мозга, играющие специфицированные роли, начинают словно бы интенсивно перезваниваться (анатомия мозга такова, что интеграция такого рода возможна, полагают они, лишь в таламо-кортикальной области и интегрированным может оказаться лишь небольшое число процессов: поэтому содержание непосредственных актов сознания ограничено) — такое коммуницирование и составляет фундамент сознания. Эдельмана считают образцовым представителем холистского подхода к субъективности. Противоположностью такого понимания является гипотеза, выдвинутая Р. Пенроузом и С. Хамероффом. Они утверждали, что сознание может быть результатом квантовых эффектов в микротрубочках нейронов⁵⁰.

К некоторым из упомянутых гипотез мы еще вернемся. Пока же ограничимся сказанным. Надеюсь, читатель или читательница почувствовали, что исследования сознания в наши дни притягивают к себе множество интереснейших людей из самых разных областей. Это такой интеллектуальный водоворот, в котором кружатся сотни идей, научных концепций и даже художественных образов — от *The Matrix* до *La misteriosa fiamma della regina Loana* в недавнем романе У. Эко. Но вот вопрос, ответ на который неочевиден: что делать в нем философам? Суть в том, что большинство идей, которые были кратко рассмотрены выше, не имеют прямого отношения к философии. Определение нейронных коррелятов сознания, моделирование функций различных ментальных модулей, уточнение числа и характеристик этих модулей на

основе психологических экспериментов — все это вопросы, относящиеся к ведению экспериментальной науки.

Можно, разумеется, повторить, что философы должны налаживать диалог между разными областями науки, имеющими выход на сознание, -- между нейробиологами и ког-нитивистами, эволюционистами и психологами, между лингвистами и квантовыми физиками — сочетания могут быть самыми необычными. Кроме того, они могут снабжать ученых разного рода методологическими принципами, задавать хорошие вопросы (ответ на которые, возможно, не их дело)⁵¹ и прояснять понятия, использующиеся при изучении сознания.

Все это, конечно, так, и очень многие ученые в наши дни признают необходимость философии для решения подобных задач. Но сами философы не всегда готовы ограничиваться ролью служанок науки. Им хочется верить, что в сфере исследования сознания есть такие вопросы, на которые может дать ответ именно философия⁵².

Одной из тем, на обсуждение которых по праву могут претендовать философы, является та, что вынесена в заглавие этой книги. Это так называемая «трудная проблема сознания» (the hard problem of consciousness). Выйти к ней проще всего, отталкиваясь от поисков нейронных коррелятов сознания, о которых мы только что говорили. Допустим, мы выяснили, что сознание тех или иных данностей у некоего человека сопровождается активностью определенных коалиций нейронов. В таком случае многие скажут, что мы обнаружили нейронные корреляты сознания. Но можно задать вопрос: а почему эта активность нейронов вообще сопровождается сознанием? Почему дело не ограничивается одной этой активностью, имеющей определенные последствия для поведения организма в целом? Ведь кроме нее есть еще что-то — переживание таких процессов.

Это переживание (природу которого удачно выразил Т. Нагель формулой «каково это быть» тем-то или тем-то)⁵³ мы и называем сознанием. Но кажется, что оно — это переживание — совсем не необходимо, кажется, что его могло бы и не быть. И все же оно есть, с этим надо как-то считаться и, самое главное, его надо как-то объяснять. А именно: надо объяснять, во-первых, почему оно вообще существует и, во-вторых, каким образом это переживание соотносится с его нейронными коррелятами. В такой формулировке, впрочем, уже предполагается ответ: отношение сознания и нейронных процессов может быть отношением корреляции. При таком отношении имеются две реальности, онтологически независимые, но сопутствующие друг другу. Однако, поскольку между со-

знанием и нейронными процессами могут быть и другие отношения, совсем не обязательно предвосхищать окончательное решение и формулировать обозначенную выше проблему именно в таком виде. Можно выразить ее и иначе: каким образом сознание как специфический феномен относится к соответствующим ему нейронным процессам?

Эта формулировка лучше, но и она не является безупречной — главным образом потому, что из нее неясно, в каком смысле в ней используется понятие сознания. Напомню, что о сознании можно говорить по меньшей мере в двух аспектах. В узком смысле сознание является чем-то, что противоположно бессознательному⁵⁴. Не исключено, что такого рода сознание не является специфическим феноменом и что главное в нем — функциональная организация, которая может быть исчерпывающе описана на нейронном или на более абстрактном функциональном уровне. Поэтому в контексте «трудной проблемы» лучше говорить о сознании в широком смысле, т. е. о ментальных состояниях, будь то неинтенциональных, таких, как некоторые удовольствия, или интенциональных, вроде пропозициональных установок — убеждений, желаний и т. п. Здесь непросто отрицать, что каждое из подобных состояний имеет специфическую качественную сторону (поэтому их часто так и называют — *qualia*⁵⁵, т. е. «качества»)⁵⁶, которую сложно свести к функциональным параметрам и которая очевидно отлична от данности нейронных процессов.

Так что «трудную проблему» лучше всего сформулировать следующим образом: как ментальные состояния, в их качественном аспекте, соотносятся с соответствующими им нейронными процессами и почему они вообще существуют в привязке к этим процессам?

Обозначенные выше вопросы едва ли могут быть исчерпаны эмпирическими методами. Конечно, опыт может убедить нас в истинности общего тезиса о том, что мозг порождает сознание. Но это не снимет «трудную проблему», а лишь конкретизирует ее: как и почему мозг порождает сознание? Более детальные данные в лучшем случае помогут найти устойчивые связи между нейронными процессами и ментальными состояниями — или, наоборот, показать, что их не существует. Подобные результаты могут стать основой для тех или иных гипотез об их соотносительном онтологическом статусе.

Но для полноценных ответов нужны не гипотезы, а однозначные схемы, возможные лишь при демонстрации концептуальной необходимости. Однако, чтобы продемонстрировать ее, недостаточно использовать стандартные аналитические методы прояснения значений тех или иных слов⁵⁷: надо вдумываться в сам способ нашего представления ментального и

и физического, выяснять, какие априорные принципы включены в представление о физическом и по каким принципам мы приписываем физическим объектам ментальные состояния. Мы должны также смотреть, как сочетаются эти принципы. Это позволит уточнить, каким образом мы должны мыслить отношение ментальных явлений и физических процессов в мозге. Те, кто прочел введение к данной книге, уже знают, что, на мой взгляд, именно такие методы характеризуют научную философию. Поэтому «трудная проблема сознания» — собственно философский вопрос. Законность его мало кто оспаривает⁵⁸. Но вот можно ли найти на него ответ? Это тоже вопрос. Ему-то как раз и посвящена эта книга.

Внимательные читатели, конечно же, заметили, что я все время закавычиваю словосочетание «трудная проблема сознания». Так происходит не потому, что я внутренне не согласен с этим выражением, а потому, что это, собственно, термин, введенный совсем недавно. Он был придуман австралийским философом Дэвидом Чалмерсом и презентован им на конференции «К научной основе сознания», которая состоялась в Туссане (Аризона, США) в 1994 г. (это был первый опыт, с тех пор подобные конференции — под новым названием «На пути к науке о сознании» - проводятся в Туссане каждые два года и, как правило, вызывают горячий интерес всего «тесного мира»)⁵⁹.

Конференция 1994 г., прошедшая в этом царстве кактусов, громадных моллов и бесплатных автобусов, курсирующих по пустынному центру города, стала настоящей легендой — именно из-за презентации на ней термина «трудная проблема сознания», растиражированного с тех пор миллионами копий. Вот как рассказывает об этом событии Стюарт Хамерофф, нынешний директор Центра исследований сознания в Туссане: «В день открытия первым выступал очень известный, знаменитый философ, но это была ужасно скучная речь, второй докладчик тоже не блистал, и я уж начал волноваться — ну, вроде драматурга в день премьеры, — что все кончится провалом⁶⁰. Третьим был неизвестный молодой философ по имени Дэвид Чалмерс, с волосами до пояса, в футболке и джинсах, и это был лучший из всех докладов, которые я когда-либо слышал о сознании. Он рассуждал о легких проблемах сознания (таких, как возможность дать отчет, восприятие и т. п.), а затем также и о трудной проблеме сознательного опыта, о том, "каково это быть", о квалиа, или сырых ощущениях. Затем была кофейная пауза и — будучи одним из организаторов конференции — я ходил среди участников и прислушивался, словно драматург в день премьеры. И все они жужжали только о докладе Дэйва и о его "трудной проб-

леме". Думаю, этот момент действительно гальванизировал международное движение по изучению сознания, поскольку проблема была найдена»⁶¹.

Мы видим, что «трудная проблема сознания» изначально была противопоставлена Чалмерсом «легким проблемам». Именно о «легких проблемах» шла речь в начале этой главы. Эти проблемы «легки» не в том смысле, что их решение очевидно, а потому, что в принципе понятно, как их решать (по крайней мере так кажется на первый взгляд), а именно: стандартными методами объективных эмпирических наук — моделированием, экспериментами и т. п. Обобщая, можно сказать, что «легкие проблемы сознания» отвечают на вопрос «как работает сознание?», тогда как решение «трудной проблемы» предполагает уточнение онтологического статуса сознания.

Уточнение онтологического статуса сознания подразумевает объяснение отношения сознания и мозга, или ментального и физического. И в этой связи можно усомниться в новизне такой постановки вопроса. Философы тысячи лет пытались ответить на него. Поиски ответа особенно интенсифицировались в Новое время, когда Декарт провел резкую границу между мыслящими и протяженными вещами⁶² и этим инициировал многочисленные дискуссии о том, каким образом такие вещи соотносятся друг с другом. В XVII и XVIII веках были перепробованы, кажется, все возможные решения. Не забывали эту проблему и в последующие времена. Особенно актуальной — под именем «проблема сознание—тело» (mind-body problem) — она стала в XX веке, задолго до выступления Чалмерса в Туссоне. Почему же тогда ее новое терминологическое одеяние (хотя вскоре мы увидим, что дело не только в нем) произвело столь мощный эффект? Чалмерс предлагает ответ на этот вопрос. Он считает, что термин «проблема сознание—тело» как бы размыт и скрывает в себе несколько разных вопросов — что мешает сфокусироваться на сути.

Речь идет о следующих вопросах: как мозг может быть основой субъективного опыта? как мозг может быть основой мышления, или рациональности и интеллекта? как сознание может оказывать влияние на физический мир? Все они, по мнению Чалмерса, «тесно связаны друг с другом, но все же являются несколько разными проблемами»⁶³. К примеру, если допустить, что мышление — это некая алгоритмическая деятельность, то вопрос, как мозг может быть основой мышления, сводится к выяснению, каким образом в нем инсталлированы те или иные алгоритмы. И этот вопрос может быть решен без обращения к проблеме онтологического статуса сознания. Третий из упомянутых Чалмерсом вопросов — о вли-

янии сознания на тело — на первый взгляд не связан с уточнением механизма порождения сознания, хотя позже мы увидим, что именно решение этого вопроса позволяет прояснить данный механизм. «Трудная проблема», тем самым, концентрируется вокруг первого вопроса: «Как мозг может быть основой субъективного опыта?» И именно он, по Чалмерсу, является «реальным ядром проблемы сознание—тело»⁶⁴.

Хочу обратить внимание, что процитированные места взяты из интервью с Чалмерсом, проведенного С. Блэкмор. В этом интервью, помимо прочего, Чалмерс утверждает, что его «трудная проблема» не является чем-то новым⁶⁵. И в приведенной выше формулировке действительно слышатся традиционные мотивы, если истолковать ее как вопрос о том, каким образом сознание может быть встроено в физическую реальность. Но дело в том, что в оригинальном, «туссанском» виде «трудная проблема сознания» была поставлена несколько иначе (и в ее предварительном обсуждении несколькими страницами выше я учел это), вбирая не только вопрос о том, как объяснить зависимость сознания, или, как иногда выражается Чалмерс, «опыта» (experience), от мозга, но и о том, почему мозг является его основой: «Многие согласны, что опыт возникает на физической основе, но мы не можем как следует объяснить, почему и как он возникает подобным образом»⁶⁶. Этот аспект «почему» не имеет явных прецедентов в прошлом, но сейчас мы сосредоточимся на аспекте «как», тем более что они, конечно же, связаны. Акцентировка на этом моменте, т. е. на положении, что существует сознание и мозг, и задача состоит в том, чтобы понять, как можно мыслить отношение этих реальностей и зависимость одного от другого полезна именно тем, что она позволяет разместить этот вопрос в традиционные контексты и рассматривать ответы на него даже со стороны мыслителей, не принимающих терминологию «трудной проблемы», но занимающихся сходными вещами. И прежде всего надо заметить, что, даже если согласиться, что Чалмерсу удалось сфокусировать проблему сознания, это еще не должно означать, что она будет решаемой. Далеко не все разделяют оптимизм по этому поводу. В конце XX века появилась даже целое направление «новых мистерианцев» (термин О. Фланагана⁶⁷), т. е. философов, утверждающих, что проблема отношения сознания и мозга должна остаться «тайной» (mystery).

Во главе этого направления — американский философ с необъятными интересами (в поле которых попадают и Шекспир, и формальная логика, и кинематограф, и этика, и мета-философия) Колин Макгинн⁶⁸, но его вдохновителем является Н. Хомский, который в свое время провел различие между

«проблемами», допускающими ясное решение, и «тайнами», в понимании которых нет никакого прогресса⁶⁹. Надо, правда, сказать, что сам Хомский не спешит объявлять проблему сознания тайной. К числу таковых он относит, скорее, вопрос о внутренних причинах поведения⁷⁰. Но его последователи -Макгинн, а также зависящий от Макгинна в этом пункте С. Линкер — обобщили разрозненные замечания Хомского и составили целый реестр тайн, включающий отношение сознания и мозга⁷¹.

Аргументы, приводимые Макгинном в пользу своего тезиса о таинственности отношения сознания и мозга — он утверждает, что, хотя мы знаем, что сознание зависит от мозга, мы в принципе не можем понять природу этой зависимости⁷², - базируются на уже знакомой нам модульной теории сознания. Если сознание — это совокупность высокоспециализированных механизмов, каждый из которых настроен естественным отбором на реализацию вполне определенной функции, то вполне может оказаться, что имеется некий класс самих по себе несложных вопросов, которые, однако, являются недоступными для решения именно в силу отсутствия в нашей ментальной системе модулей, предназначенных для ответа на них⁷³.

Если такие вопросы действительно существуют, то нам придется признать, утверждает Макгинн, что мы «когнитивно замкнуты» по отношению к ним — подобно тому, как животные когнитивно замкнуты относительно проблем, которые решает эмпирическая наука. Мы не замкнуты по отношению к ним — ведь это мы практикуем такую эмпирическую науку, - но замкнуты по отношению к другим проблемам. И в таком случае это уже не проблемы, а тайны.

Общая программа Макгинна вполне понятна, и остается лишь уяснить, по каким критериям мы можем установить свою замкнутость относительно тех или иных познавательных вопросов. Здесь аргументация Макгинна не столь прозрачна, и он порой модифицировал ее. Иногда — к примеру в трактате «Таинственный огонь» (1999) — он высказывался в том ключе, что человеческое познание распространяется лишь на те области, где оно может работать с комбинациями каких-либо исходных элементов. Например, мы в состоянии понять и объяснить химические свойства вещества, исходя из особенностей атомов, составляющих его молекулы. Но ментальные состояния, такие, как радость или боль, не могут быть объяснены как результат сочетания тех или иных нейронных процессов. Именно поэтому их отношение к нейронным процессам непостижимо⁷⁴.

Этот аргумент не выглядит безупречным. В самом деле, даже если согласиться с комбинаторной интерпретацией че-

ловеческого интеллекта, которую предлагает Макгинн, из нее не следует, что анализ некоей материальной системы не может привести к выводу, что ее функционирование нельзя объяснить исходя из характеристик образующих ее физических компонентов. В таком случае пришлось бы допускать существование какого-то дополнительного компонента, качественно отличного от других компонентов системы. Именно этот дополнительный компонент позволял бы ликвидировать бреши в объяснении. Но если этот компонент — не гипотетический конструкт, а нечто наличное во внутреннем опыте, то тогда, наоборот, вся эта схема может объяснять это ментальное состояние, т. е. она может показывать, почему такое состояние должно естественным образом обнаруживаться при данном характере физических процессов.

Таким образом, ничто не мешает объяснять существование определенных параметров материальных систем, даже если они не могут быть истолкованы как результат сочетания ее физических компонентов. Поэтому более перспективной кажется другая линия «мистерианских» рассуждений Макгинна, а именно та, которая сводится к тезису, что ментальные и физические состояния открываются людям, соответственно, во внутреннем и внешнем чувствах⁷⁵, но ни то ни другое не раскрывает нам «ноуменальное» свойство, от которого зависит связь мозга и сознания⁷⁶. И мы никогда не узнаем его, так как, кроме указанных чувств, у нас нет выходов к реальности.

Эта позиция напоминает о кантовском различии вещей в себе и явлений и о кантовском же положении о непознаваемости первых. Сходство не случайно — сам Макгинн ссылается на Канта⁷⁷. Упоминание Канта в таком контексте может, однако, и не ограничиваться ссылкой на его концепцию вещей в себе и явлений. Суть в том, что Кант предлагал в чем-то похожее на макгинновское решение проблемы отношения сознания и мозга — хотя кантовское решение было еще более радикальным и скептическим. И это скептическое решение состояло в том, что человеческий разум не в силах установить, каким образом они относятся друг к другу⁷⁸. Этот вопрос выходит за пределы человеческого познания. Все, что мы можем сделать — констатировать эмпирическую корреляцию ментальных и физических феноменов, данных нам во внутреннем и внешнем чувствах.

Неудивительно, что некоторые авторы настойчиво указывают на актуальность суждений Канта по данному вопросу⁷⁹ — в том же Туссане не раз слышались призывы обратиться к кенигсбергскому мыслителю. Но тут возникает вопрос: а почему только к Канту? Кантовский скептицизм относительно возможности решения проблемы онтологического ста-

туса сознания звучал уже в «Опыте о человеческом познании» Локка, да и в ряде других работ той эпохи. Вообще, как я уже отмечал, в XVII и особенно в XVIII веке были опробованы чуть ли не все возможные подходы к данному вопросу. Поэтому нас не должно удивлять то обстоятельство, что периодически раздаются фразы о том, что современная философия сознания изобретает велосипед, «открывая» уже изрядно протоптанные пути.

Эти фразы звучат в исполнении весьма влиятельных людей. На эту тему, к примеру, любит порассуждать Хомский. Еще более популярны такие пассажи среди российских историков мысли⁸⁰.

Надо признаться, что и я поддавался этому искушению и в своей книге о философии сознания эпохи Просвещения даже высказал идею, что эта философия может «просветить» современных авторов, указывая им действительно перспективные направления исследований и в то же время отсекая нерешаемые вопросы, к числу которых, как я думал, относится и наша «трудная проблема».

Сейчас я придерживаюсь иного мнения. Доводы Макгин-на и его сторонников не убеждают меня хотя бы потому, что они даются им, как, впрочем, и Канту, слишком дорогой ценой, а именно ценой введения понятия реальности самой по себе, принципиально отличной от эмпирических данностей. Хорошо известно, насколько опасным является этот шаг, не только чреватый концептуальными противоречиями, но и бросающий вызов здравому смыслу, для которого реальностью является именно мир чувственных предметов⁸¹.

Так что я больше не считаю «трудную проблему сознания» неразрешимой. Кроме того, я полагаю, что современные контексты дискуссий о ней существенно отличаются от тех, которые можно обнаружить в эпоху Просвещения. Наконец, мне кажется, что в трактовке этой проблемы в наши дни имеет место реальное движение вперед. Все это, разумеется, надо обосновывать. Этим я и собираюсь заняться в ближайшее время. В конце данной главы я укажу на самые заметные отличия современных трактовок проблемы сознания от тех, что имели место в Новое время, и уточню, в чем именно, на мой взгляд, состоит тот прогресс, который был уже достигнут. В трех последующих главах мы рассмотрим наиболее значительные теоретические построения, воздвигнутые за последнее время в ходе нового осмысления онтологического статуса сознания. Мы увидим, что, несмотря на их многочисленные достоинства, они не выводят нас из концептуальных тупиков. После этого мы обсудим еще одно решение, которое, надеюсь, вызовет нас из лабиринта.

Итак, посмотрим прежде всего на различия контекстов. Чтобы прояснить их, задумаемся о том, какие вообще мыслимы трактовки статуса ментального и физического. Во-первых, есть крайние варианты: не существует либо ментального, либо физического. Далее, имеется среднее решение: и ментальное, и физическое реальны. Но тогда возникает вопрос о том, как они относятся друг к другу. Либо они онтологически самостоятельны, но в этом случае надо объяснить, почему ментальные состояния соответствуют физическим процессам, либо одна из этих реальностей производна от другой, либо, наконец, обе эти реальности — свойства третьей, не данной в опыте. Все эти варианты — дополняющие скептический ответ — можно специфицировать, но каркас решений составляют именно они.

Так вот, в современной философской литературе доминирует один из перечисленных вариантов, а именно тот, согласно которому ментальное производно от активности мозга как физической системы.

Совершенно другая картина наблюдалась в эпоху Просвещения (я выделяю ее потому, что в XIX веке эта тема ушла на второй план, а в более ранние времена в ее трактовке преобладали архаичные подходы)⁸². Определенный вес имели «идеалистические» системы, авторы которых отрицали существование физической реальности. Речь идет прежде всего о системе Беркли, но были и другие. В целом они, однако, вызвали отторжение философского сообщества. Наибольшее распространение в конце XVII — начале XVIII века получили теории, признававшие независимый онтологический статус (субстанциальность) ментального и физического, или души и тела.

Но при таком понимании, как уже отмечалось, надо было решать вопрос о возможности соответствия физических и ментальных состояний — то, что такое соответствие имеет место, принималось как факт. При рассмотрении этого вопроса был выдвинут целый набор конкурирующих теорий. Самая простая из них — теория естественного⁸³ влияния. Согласно этой теории, душа и тело способны напрямую взаимодействовать, несмотря на различие их природ. К этому решению склонялся Декарт, основатель новоевропейской философии сознания⁸⁴. Однако многие его последователи считали, что различие природ души и тела⁸⁵ является достаточным основанием для отрицания возможности их прямого взаимодействия⁸⁶. Отсюда они выводили, что соответствие ментальных и психических состояний должно иметь своим источником некую третью сущность, возвышающуюся над сферами физического и ментального, а именно Бога.

Бог может координировать ментальные и физические изменения либо как их безличная основа, в которой они неразличимы — это позиция Спинозы, — либо непосредственным участием в каждом конкретном случае — Мальбранш и другие окказионалисты, — либо через упорядочение ментальных и физических процессов, произведенное им раз и на все времена еще при творении мира — теория «предустановленной гармонии между душой и телом» Лейбница и Вольфа.

Сторонники всех этих теорий остро соперничали друг с другом. В динамике это выглядело следующим образом. Декартовская теория естественного влияния сменилась окказионалистскими подходами, которые, однако, имели большие недостатки, так как допускали ситуацию постоянного чуда, неизбежного при повсеместном вмешательстве Бога в физические и ментальные процессы. Этот дефект предопределил вытеснение окказионализма теорией предустановленной гармонии Лейбница, являвшей собой антропоморфизированный вариант спинозовского параллелизма. Лейбницевская концепция была подхвачена Вольфом, а затем и сотнями его последователей.

Между тем в середине XVIII века стало очевидно, что лейбницевольфовская система предустановленной гармонии души и тела плохо согласуется с общей идеологией набирающего силу Просвещения. Эта идеология опиралась на здравый смысл, для которого учение о предустановленной гармонии казалось слишком уж экстравагантным и фантастичным. Поэтому с 40-х гг. XVIII столетия все больший вес обретает «исправленная», или «улучшенная», теория естественного влияния⁸⁷. Дальше — больше. Европейские мыслители начинают все чаще говорить о том, что если физическое может воздействовать на ментальное, то почему бы не предположить, что оно может порождать ментальное? При таком подходе сразу отпадает вопрос о возможности взаимодействия двух разнородных субстанций — отпадает просто потому, что двух субстанций не существует. Есть одна субстанция — материя, и она продуцирует сознание.

Возможность такого решения в принципиальном плане обосновывал еще Локк, но он не считал его единственным вариантом. Идеи Локка радикализировал Вольтер, провозгласивший сознание свойством мозга. Тезис Вольтера подхватили и развили его более молодые соотечественники — Ламетри, Дидро и Гольбах. Получили они прописку и за пределами Франции. В Германии подобные взгляды отстаивал, к примеру, Хисман, в Англии — Пристли⁸⁸. Таким образом, натуралистические решения, характерные для современной дискуссионной среды, сформировались уже в те времена.

Но они, еще раз подчеркну, не были доминирующими. В XVII и XVIII веках интересующая нас проблема сознания рассматривалась главным образом в дуалистической и теологической перспективе, которая совершенно маргинальна в наши дни⁸⁹. В этом — существенное отличие старых и новых контекстов «трудной проблемы».

И сразу отмечу, что сам этот факт является свидетельством, позитивных изменений. Ведь теологическая заряженность большинства прежних систем - как убедительно показал еще Кант — это признак их глубинной догматичности, которая противоречит самой природе науки и философии. При выведении же за скобки понятия Бога как объяснительного принципа количество возможных решений вопроса об онтологическом статусе сознания резко сужается. И самым предпочтительным может оказаться именно тот путь, который избирает большинство современных авторов и который в свое время уже опробовали Ламетри, Пристли и другие радикальные умы.

Исходя из сказанного, можно предположить, что сегодняшняя философия сознания является наследницей психологического натурализма Просвещения — с оговоркой о разрыве цепи преемственности, вызванного возрождением теолого-субстанциалистских моделей в конце XVIII и XIX веке как в немецкой философии у Фихте, Шеллинга и Гегеля, так и в других традициях — от Мен де Бирана и Кузена до Брэдли. Однако, хотя в этом тезисе о косвенном наследовании есть рациональное начало, сразу надо предостеречь от его неверных толкований. Суть в том, что при некотором внешнем сходстве с психологическими натуралистами эпохи Просвещения современные философы сознания решают абсолютно другие задачи. И понимание этого обстоятельства, как представляется, должно окончательно убедить читателей этой книги в позитивном отличии современных контекстов «трудной проблемы» от того, что было раньше.

Новые задачи, о которых я сказал, состоят в следующем: если натуралисты XVIII века пытались доказать, что сознание есть продукт мозга, и привлекали для этого как факты, говорящие о зависимости ментальных состояний от физиологических процессов — классическая работа на эту тему - это, конечно, «Человек-машина» (1747) Ламетри, — так и разного рода доводы, вроде того, что ментальное не противостоит физическому, так как ему тоже свойственна некая протяженность (здесь особенно показательны рассуждения Пристли), то современные авторы исходят из тезиса, что мозг так или иначе порождает сознание, и хотят прояснить это отношение.

В этом все дело. Аналитические философы осознали недостаточность простых констатаций производности сознания от мозга. Это не конец исследования природы сознания, а, скорее, только начало.

На мой взгляд, речь идет об очень важных вещах. Поэтому еще немного задержусь на них. Итак, мы имеем нечто вроде консенсуса в отношении того, что мозг порождает сознание (хотя и это не догма, а просто самое естественное объяснение факта динамического соответствия физических и ментальных процессов). Но ведь это, как сказали бы логики, весьма «слабый» тезис — и не случайно я использую для его формулировки метафорическое выражение «порождает». Данный тезис никак не специфицирует ответ на вопросы: как мозг порождает сознание и почему он порождает его? От ответа на них — т. е. от решения «трудной проблемы» — зависит уточнение онтологического статуса сознания. А поскольку ответы могут быть разными, это означает, что здесь может быть место для различных, возможно даже несовместимых теорий. В любом случае столь расплывчатая формулировка, как «мозг порождает сознание», должна подталкивать философов к тщательным изысканиям. Причем эти изыскания не так уж сильно связаны, скажем, с нейробиологическими исследованиями. Допустим, мы установили, какие нейронные процессы соотношены с конкретными ментальными состояниями. Это не снимет проблему, так как нам еще надо будет понять, почему эти процессы вызывают такие состояния и реальна ли возможность, что их будет вызывать что-то другое.

Одним словом, при прояснении формулы «мозг порождает сознание» речь должна идти не о фактическом, а о сущностном анализе их соотношения. Но чтобы произвести этот сущностный анализ, совсем не обязательно допускать у нас некую способность проникать в глубины сознания и материи. Сущностный анализ в данном случае совпадает с концептуальным анализом, в уже объясненном смысле: мы должны прояснять сами понятия ментального и физического, т. е. необходимые аспекты нашего мышления о последних, и на этом пути мы сможем выявить то отношение между ментальным и физическим, которое стыкуется с этими концептуальными схемами.

Вот, стало быть, что стоит за прояснением тезиса о зависимости сознания от мозга. Я специально сформулировал это в достаточно абстрактном виде, чтобы читатели ощутили, что за этим тезисом могут скрываться непростые проблемы. Но реальная ситуация еще более интересна — как мы увидим, эти проблемы не только не просты, здесь нечто большее: они прямоком ведут к парадоксам.

Иначе говоря: я утверждаю, что, когда мы пытаемся уточнить внешне безобидное положение о зависимости сознания от мозга, мы очень быстро обнаруживаем, что можем оказаться в тупике. Более того, возникает чувство, что, как бы мы ни шли, тупик неизбежен. Это довольно странная ситуация — в чем нам еще предстоит убедиться в ближайших главах. Ведь в них как раз и пойдет речь о возможных — и разнонаправленных — спецификациях Тезиса о порождении мозгом сознания и, как можно предположить, о разных вариантах решения «трудной проблемы сознания». Эти спецификации в известном смысле взаимодополнительны. Более того, может показаться, что они исчерпывают весь спектр возможных решений в контексте общего положения о зависимости сознания от мозга.

Одно из этих решений предложено Джоном Сёрлом, другое — Дэниелом Деннетом, третье — Дэвидом Чалмерсом. Поскольку учения этих мыслителей⁹⁰, олицетворяющих современную аналитическую философию сознания, относительно самостоятельны, порядок их рассмотрения и интерпретации по большому счету неважен. И все же одна из возможных последовательностей представляется более удачной. Объясняется это тем, что система Чалмерса может быть истолкована как своеобразный синтез идей Сёрла и Деннета. Поэтому ее лучше оставить напоследок. Слова о синтезе не могут не пробудить в памяти профессионального философа термины «тезис» и «антитезис». Деннет и Сёрл, кстати, хорошо подходят для их иллюстрации: они и впрямь неприкрытые противоположности — судя хотя бы по их взаимным обвинениям в «интеллектуальной патологии». Чью же концепцию лучше принять за «тезис»? Припомним Гегеля, трудно удержаться, чтобы не начать с того, кто предлагает внешне простое позитивное решение проблемы сознания, которое соответствует установкам здравого рассудка, исходящего из признания реальности внутреннего мира психических состояний, а на вторую, антитегическую позицию поместить философа, который, напротив, борется с обыденными установками и хочет доказать несуществование или хотя бы иллюзорность сознания. Случайно или нет, но наши будущие герои идеально вписываются в эти характеристики. Сёрл выступает апологетом здравого смысла в решении загадки ментального, а Деннет, напротив, изобретательно разрушает «народный» образ сознания. Так что начнем с Сёрла, а продолжим Деннетом, хотя, конечно, в рассказ будут включены и многие другие персонажи.

Примечания

¹ Это не значит, конечно, что периодически кто-нибудь не пытается возобновить эти темы. См., напр.: Суинберн Р. Философия религии в англо-американской традиции // Философия религии: альманах. М., 2007. С. 89-136.

² Я не разделяю точку зрения Ф. Ч. Коплстона, согласно которой историческое движение философской мысли сводится к переодеванию «вечных проблем» в новые культурные платья.

³ Дальше всех пошел, пожалуй, писатель Тибор Фишер, намекавший, что преподавателям философии в такой ситуации пора заняться чем-то более серьезным, к примеру грабежом банков.

⁴ См., напр.: Рорти Р. Философия и зеркало природы. Новосибирск, 1997. С. 264-292.

⁵ Основатели аналитической философии Фреге, Рассел, Мур и Витгенштейн не жаловали сам термин «аналитическая философия», но это не значит, что он имеет позднее происхождение. А. Престон отмечает, что в «номинативном смысле», т. е. как обозначение школы или направления, он употреблялся уже в 30-е годы XX века Э. Нагелем и Р. Коллингвудом — см.: Preston A. Prolegomena to any future history of analytic philosophy // *Metaphilosophy* 35: 4 (2004). P. 454. В описательном же смысле он встречался, по его наблюдениям, уже в конце XIX в. В действительности этот современно звучащий термин, как, к примеру, и термин «философия языка», имеет еще более глубокие корни. Так, в «Философско-медицинском словаре» Ф. Й. Циммермана «аналитическая философия» определялась как «прояснение неясностей в представлениях о наших принципах» — Zinimermann F. J. *Philosophisch-medicinisches Wörterbuch*. Wien, 1803. S. 190. С такой дефиницией не стал бы по существу спорить ни один аналитический философ наших дней. Развитие аналитической философии в XX в. было таким стремительным, что историки философии не поспевали за его осмыслением. Ясно лишь, что на рубеже XXI в. она вступила в новый, «конструктивный» этап и на первые позиции вышли авторы, имеющие склонность к созиданию систем. Прежде всего это Деннет и Сёрл. В последние десятилетия также неуклонно возрастало насыщение аналитических текстов экспериментальными данными — ср.: Searle J. R. *Freedom and Neurobiology: Reflections on Free Will, Language, and Political Power*. N. Y., 2007. P. 30-31.

⁶ См.: McGinn C *The Making of a Philosopher: My Journey through, Twentieth-Century Philosophy*. N. Y., 2003 (2002). P. 187; Dummett M. *Origins of Analytical Philosophy*. Cambridge MA, 1993. P. 2.

⁷ McGinn C. *The Making of a Philosopher*. P. 214. Аргументы, приводимые Макгинном, выглядят убедительно: он просто перечисляет ведущие университеты, расположенные в районе Нью-Йорка. Рассуждая об американской философии, следует напомнить, что она не является чем-то однородным в идейном плане (хотя философское сообщество в США скорее цельно, чем раздробленно, что объясняется в том числе и устоявшейся практикой запрета на работу защитив-

шимся аспирантам в выпускающем университете — это укрепляет межуниверситетские связи). Сами американские философы полагают (см. напр.: Marsoobian A. T., Ryder J. (eds) *The Blackwell Guide to American Philosophy*. Malden, 2004. P. XV), что в ней присутствуют три основные линии — собственно американская, т. е. прагматистская, континентальная и аналитическая. Но подавляющее большинство кафедр философии контролируют аналитики. В частной беседе Дж. Райдер утверждал, что в США есть три-четыре университета, кафедры философии которых работают в континентальной традиции, и два центра прагматизма; при общем количестве философских кафедр, по данным Американской философской ассоциации за 2002 г., более 1730 (из них около 200 с Ph.D программой; среднее количество постоянных сотрудников — 6; ну а лучшей кафедрой философии англоязычного мира несколько лет подряд признается кафедра Нью-Йоркского университета (NYU) — среди 20 ее сотрудников Т. Нагель и Н. Блок). Что касается истоков аналитического движения в США, то решающую роль в его зарождении сыграли иммигранты из Европы. А. Бекерман приводит любопытное свидетельство о мнении У. Куайна на этот счет: он считал, что «начало аналитической философии в США можно точно датировать. В 1935 г. Куайн, Гудман и другие коллеги сопровождали Рудольфа Карнапа, выступавшего перед Философской ассоциацией в Балтиморе. После доклада Карнапу пришлось отвечать на возражение Артура Лавджоя, и он сделал это в характерной для него и для аналитической философии манере: "Если Артур Лавджой имеет в виду А, то р, а если он имеет в виду В, то q"» — Beckermann A. *Analytische Einführung in die Philosophie des Geistes*. 2 Aufl. В., 2000. S. VIII.

⁸ Удачный обзор главных идей Скиннера, в том числе и философских, дан в книге Л. Стевенсона — см.: Стевенсон Л. *Десять теорий о природе человека*. М., 2004. С. 187-203.

⁹ К созданию этой книги, как вспоминал сам Скиннер, его подтолкнул один философский разговор за обедом — с небезызвестным А. Уайтхедом. Беседа состоялась еще в тридцатые годы. Соглашаясь со Скиннером относительно эффективности бихевиористских объяснений, Уайтхед заметил, что на человеческую речь распространить их все же не удастся. В подтверждение своей мысли он предложил Скиннеру бихевиористски истолковать хотя бы одну фразу, которую он тут же и произнес: «На этот стол не падает черный скорпион». Тогда-то у Скиннера и возник замысел «Вербального поведения» — см.: Skinner B. F. *Verbal Behavior*. Acton, 1957. P. 456-457.

¹⁰ Chomsky N. A Review of B. F. Skinner's *Verbal Behavior* // *Language* 35: 1 (1959). P. 26-58.

¹¹ «Вербальное поведение», кстати, Скиннер определял как «поведение, подкрепляемое при посредстве других людей» — Skinner B. F. *Verbal Behavior*. P. 14. Т. е. он понимал его в очень широком смысле. Речь, или «вокальное поведение», — лишь одна из его разновидностей.

¹² Книга Скиннера и в самом деле насыщена такими «кризисными» для бихевиористской идеологии оборотами, как «приватные стимулы», «субъективное поведение» и т. п.

¹³ См., напр.: Chomsky N. *Recent contributions to the theory of innate ideas*//*Syntheses* 17(1967). P. 2-11.

¹⁴ См. на эту тему: Хомский Н. Картезианская лингвистика. М., 2005, а также: Хамский И. О природе и языке. М., 2005. С. 136-233,

¹⁵ Разве что Дж. Фодор. Если он не шутит, то его можно легко опровергнуть — см.: Pinker S. The Stuff of Thought: Language as a Window into Human Nature. N. Y., 2007. P. 92-107.

¹⁶ Хомский отсылает к исследованиям С. Голдип-Мидоу. Ср. также факты креолизации языков.

¹⁷ Ср.: Патнэм Х. «Гипотеза врожденности» и объяснительные модели в лингвистике // Философия языка. М., 2004. С. 178-190.

¹⁸ См. об этом: Линкер С. Язык как инстинкт. М., 2004.

¹⁹ См., напр.: Chomsky N. Reflections on Language. N.Y., 1975. P. 35, 123.

²⁰ Фодор пытался связать модульную теорию сознания с незаслуженно, по его мнению, забытой «психологией способностей» — см.: Fodor J. The Modularity of Mind. Cambridge MA, 1983. P. 1-38.

²¹ См.: Fodor J, The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology. Cambridge MA, 2001 (2000). P. 5, 55-78.

²² Результаты дискуссий по модульной теории обобщил П. Кератерс. См.: Carruthers P. The Architecture of the Mind. N. Y., 2006.

²³ Блестящее изложение всей этой тематики дано в книгах С. Линкера, и прежде всего в его «Чистой доске» — Pinker S. The Blank Slate: The Modern Denial of Human Nature. N. Y., 2002.

²⁴ См., напр.: Хомский Н. О природе и языке. С. 74-75.

²⁵ Известно, кстати, что Тьюринг создавал свою Машину с оглядкой на человеческое мышление.

²⁶ На этом настаивал Р. Пенроуз в книгах «Новый ум короля» (1989) и «Тени разума» (1994). Опираясь на теорему Геделя, Пенроуз доказывал, что человеческое мышление хотя и может быть машинообразным, как правило, бывает другим - невычислимым. Однако все его аргументы были подвергнуты серьезной критике, и они мало убедили или хотя бы смутили когнитивистов.

²⁷ Putnam H. Minds and machines // Dimensions of Mind. P. 148-180.

²⁸ См., напр.: Churchland P. M. Functionalism at forty: A critical retrospective // Journal of Philosophy 102:1 (2005). P. 33-50. Черчленд критикует и другие тезисы классической когнитивной науки.

²⁹ По словам самого Патнэма, все началось с того, что он обнаружил, что компьютерная метафора плохо согласуется с его собственной экстерналистской теорией значения — см.: Putnam H. Words & Life. Cambridge MA, 1994. P. 443-444.

³⁰ См., напр, мини-статью Патнэма, присланную им на одну из конференций в МГУ — Патнэм Х. Химера когнитивной науки // Философия и будущее цивилизации: Тезисы докладов и выступлений IV Российского философского конгресса (Москва, 24-28 мая 2005 г.). М., 2005. Т. 1. С. 276.

³¹ Что, кстати, не отрицал и сам Патнэм — см.: Putnam H. Words & Life. P. 415.

³² Fodor J. A Theory of Content and Other Essays. Cambridge MA, 1990. P.198-199.

³³ См., напр.: Сакс О. Человек, который принял свою жену за шляпу. М., 2005; Feinberg T. E. *Altered Egos: How the Brain Creates the Self*. N. Y., 2001; Рамачандран В.С. Рождение разума. Загадки нашего сознания. М., 2006.

³⁴ Удачный обзор этих методик дан в книге П. Черчленда — см.: Churchland P. M. *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain*. Cambridge MA, 1996(1995). P. 151-158.

³⁵ И получив за это Нобелевскую премию в 1962 г.

³⁶ Так называемые NCC. Подробное обсуждение различных проблем и аспектов этого проекта: Chalmers D.J. What is a neural correlate of consciousness // *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, ed. by T. Metzinger. Cambridge MA, 2000. P. 17-40. См. также: Churchland P. S. *Brain-Wise: Studies in Neuro-philosophy*. Cambridge MA, 2002. P. 154-156.

³⁷ Итогом многолетних исследований стала фундаментальная книга К. Коха «В поисках сознания» — Koch C. *The Quest for Consciousness: A Neurobiological Approach*. Englewood, 2004.

³⁸ Crick F., Koch C *Towards a neurobiological theory of consciousness* // *The Nature of Consciousness: Philosophical Debates*, ed. by N. Block, O. Flanagan, and G. Guzeldere. Cambridge MA, 1997. P. 277-292. Первоначально — в *Seminars in the Neurosciences 2* (1990). См. также: Crick F. *The Astonishing Hypothesis: The Scientific Search for the Soul*. N. Y., 1995 (1994). P. 244.

³⁹ Koch C. *The Quest for Consciousness*. P. 46.

⁴⁰ Которые, однако, как считают некоторые, могут быть преодолены — см.: Hameroff S. *The good, the bad and the octopus* // *Journal of Consciousness Studies* 14: 8 (2007). P. 105-109. Ср.: Singer W. *Large-scale temporal coordination of cortical activity as a prerequisite for conscious experience* // *The Blackwell Companion to Consciousness*, ed. by M. Velmans and S. Schneider. Maiden, 2007. P. 605-615.

⁴¹ Обсуждение различных аспектов этих идей на новом экспериментальном материале — см.: Melloni L. et al. *Synchronization of neural activity across cortical areas correlates with conscious perception* // *The Journal of Neuroscience* 27: 11 (2007). P. 2858-2865.

⁴² Koch C. *The Quest for Consciousness*. P. 98, 311.

⁴³ См.: Vaars B. *A Cognitive Theory of Consciousness*. Cambridge, 1988. Новая версия его теории представлена в книге: Vaars B, *In the Theater of Consciousness: The Workspace of the Mind*. N. Y., 1997.

⁴⁴ В «Когнитивной теории сознания» понятие глобального рабочего пространства вводится Баарсом следующим образом: «Глобальное рабочее пространство есть некая информационная биржа, позволяющая специализированным бессознательным процессорам в нервной системе взаимодействовать между собой. Если искать аналог в человеческом сообществе, то им окажется школьная доска или телецентр» — Vaars B. *A Cognitive Theory of Consciousness*. P. 74.

⁴⁵ О трудностях ответа на него — см.: Churchland P. S. *Brain-Wise*. P.158-164.

⁴⁶ См.: Koch C. *The Quest for Consciousness*. P. 307, 309-310.

⁴⁷ Dehaene S., Naccache L. *Toward a cognitive neuroscience of consciousness: basic evidence and a workspace framework* // *Cognition* 79 (2001). P. 1-37.

⁴⁸ Он дает понять это в интервью с С. Блэкмор — см: Blackmore S. *Conversations on Consciousness: What the Best Minds Think about the Brain, Free Will, and What It Means to Be Human*. N. Y., 2006. P. 14, 16. С другой стороны, в более поздней статье он указывает на определенные нестыковки модели Эдельмана и Тонони с некоторыми аспектами своей теории — см.: Baars B. *The global workspace theory of consciousness // The Blackwell Companion to Consciousness*. P. 243 (на нестыковки указывали, кстати, и Эдельман с Тонони). Речь, в частности, идет о том, что метафора «трансляции», которую он использовал для характеристики глобального рабочего пространства, плохо сочетается с анатомией ряда участков мозга. Баарс готов уступить и даже допускает возможность того, что в мозге существует несколько подобных пространств — *ibid.*

⁴⁹ См., напр.: Edelman G. *Wider than the Sky: The Phenomenal Gift of Consciousness*. L., 2004; Edelman G., Tononi G. *A Universe of Consciousness: How Matter Becomes Imagination*. N. Y., 2000. Сходную теорию с середины 70-х гг. XX в. разрабатывал А. М. Иваницкий. См.: Иваницкий А.М. *Мозговая основа субъективных переживаний: гипотеза информационного синтеза // Журнал высшей нервной деятельности им. И.П. Павлова* 46: 2 (1996). С. 241-252.

⁵⁰ Hameroff S., Penrose R. *Orchestrated objective reduction of quantum coherence in brain microtubules: the "Orch OR" model for consciousness // Toward a Science of Consciousness*, ed. by S. Hameroff, A. Kaszniak and A. Scott. Cambridge MA, 1996. P. 507-540; Пен-роуз Р. *Тени разума*. Ч. 2. М., 2005. С. 206--269. Впрочем, авторы этой гипотезы полагают, что она может объяснять и единство сознания — если допустить возможность масштабной межнейронной когеренции.

⁵¹ «Философский метод проявляет свои лучшие стороны при постановке вопросов, но с ним едва ли можно преуспеть при ответах на них» — см.: Koch S, *The Quest for Consciousness*. P. 7. Нечто похожее говорил и коллега Коха Ф. Крик: «По сути философы часто задают хорошие вопросы, но им недостает технологий для получения ответов на них. Поэтому можно не обращать особого внимания на их дискуссии» — Blackmore S. *Conversations on consciousness*. P. 74. Некоторые аналитические философы, к примеру Деннет и Сёрл, не стали бы принципиально спорить с тем, что дело философов — задавать вопросы, дело ученых — отвечать на них. Подобные суждения высказывал еще Рассел в знаменитой работе «Проблемы философии» (1912).

⁵² И если бы они считали иначе, они не смогли бы утверждать, что в паши дни «первой философией» является именно философия сознания — см.: Searie J.R. *Mind, Language, and Society: Philosophy in the Real World*. N.Y., 1999 (1998). P. IX. О «центральной» философии сознания см. также: McGinn C. *Character of Mind: An Introduction to the Philosophy of Mind*. New ed. N. Y., 1996. P. 163.

⁵³ В классической статье «Каково это быть летучей мышью» -Naget T. *What is it like to be a bat? // Philosophical Review* 83: 4 (1974). P. 435-450. Впрочем, сам Нагель напоминает, что в сходном контексте этим выражением ранее пользовался Т. Спригг. А вопрос «Каково это быть летучей мышью?» задавал еще Б. Ферелл в 1950 г. См.: Nagel T. *The View From Nowhere*. N. Y., 1986. P. 15.

⁵⁴ Если бы кто-то потребовал дать более точное определение сознания, я не дал бы его. В этом нет особой необходимости. Я не знаю дефиниций ложки, но умею пользоваться ею. Так же обстоит дело и с сознанием — на первых порах вполне достаточно «рабочих» определений.

⁵⁵ Считается, что этот термин введен К. И. Льюисом в книге «Сознание и мировой порядок» (1929).

⁵⁶ Я полностью поддерживаю Сёрла, Чалмерса и других авторов, утверждающих, что все наши ментальные состояния имеют подобный «качественный» или феноменальный аспект. Это следует хотя бы из того, что, в случае их осознанности, они должны осознаваться так-то и так-то.

⁵⁷ Неважно ведь, знаете ли вы названия тех чувств, которые сейчас испытываете, или не знаете их, все равно вы испытываете их и отличаете их от данностей восприятия, имеющихся у вас, скажем, при наблюдении за работой собственного мозга, которое, возможно, и вызвало у вас те чувства. Это означает, что «трудная проблема» существует независимо от семантических контекстов.

⁵⁸ Хотя есть исключения — см., напр.: Dennett D. C. Facing backwards on the problem of consciousness; Charchland P. S. The horn-swogge problem // Explaining Consciousness — The "Hard Problem", ed. by J. Shear. Cambridge MA, 1997. P. 33-36; 37-44.

⁵⁹ Некоторый спад активности участников — если судить по количеству поданных тезисов — в 2002 и 2004 гг. (см.: Whitehead Ch. Everything I believe might be a delusion. Whoa! Tucson 2004: Ten years on, and are we any nearer to a science of consciousness? // Journal of Consciousness Studies 11: 12 (2004) был преодолен в 2006 г., и, возможно, новая волна превзойдет прежние.

⁶⁰ Хамерофф по понятным причинам не называет имен. Но из надежных источников мне известно, что первым выступал влиятельный в 90-е годы философ О. Фланаган, вторым — А. Голдман.

⁶¹ Blackmore S. Conversations on consciousness. P. 116.

⁶² Трактовка Декарта как последовательного дуалиста, впрочем, нередко оспаривается. См., напр.: Clarke D. Descartes's Theory of Mind. Oxford, 2003.

⁶³ Blackmore S. Conversations on consciousness. P. 41-42.

⁶⁴ Ibid.

⁶⁵ Ibid. P. 41.

⁶⁶ Chalmers D, J. Facing up to the problem of consciousness // Explaining Consciousness. P. 11.

⁶⁷ См.: Flanagan O. Consciousness Reconsidered. Cambridge MA, 1992. P. 109. Фланаган говорит именно о «новых мистерианцах», чтобы подчеркнуть натуралистические предпосылки их скептических построений, отсутствовавшие у их предшественников — *ibid.*

⁶⁸ Макгинн родился в Великобритании в 1950 г., учился в Манчестере и Оксфорде, работал в Лондоне и Оксфорде и перебрался в США лишь на рубеже 90-х гг.: «К концу 80-х годов философия в Британии стала депрессивным занятием» -- McGinn C. The Making of a Philosopher. P. 187.

⁶⁹ Chomsky N. Reflections on Language. P. 137-138.

⁷⁰ Позиция Хомского по проблеме сознание—тело (в ее «трудном» аспекте) весьма своеобразна. Он утверждает, что такой пробле-

мы попросту не существует, так как понятие тела с НЬЮТОНОВСКИХ времен лишено какой-либо определенности (см.: Chomsky N. *New Horizons in the Study of Language and Mind*. Cambridge, 2000. P. 84). Это звучит очень радикально, НО в действительности Хомский вовсе не отрицает существование другой проблемы — сознание—мозг. При этом он не исключает, что эта проблема будет решена вследствие прогресса нейронауки.

⁷¹ См.: McGinn C. *Problems in Philosophy: The Limits of Inquiry*. Maiden, 1993; Pinker S. *How the Mind Works*. L., 1999 (1997). P.558-565.

⁷² В автобиографии Макгини рассказывает, как он пришел к своей теории: «В моей голове перемешивались тогда два ряда мыслей: о потенциальной непознаваемости реальности и о глубоко загадочной природе отношения между сознанием и мозгом... И вот как-то ночью, когда, лежа в постели, я прокручивал в уме эти вощи... два этих ряда сомкнулись. Это было одно из тех озарений, о которых я читал в воспоминаниях других людей. Может быть, решение проблемы сознание—тело доставляет нам столько трудностей потому, что реальность содержит некий ингредиент, о котором мы ничего не в состоянии узнать. Мы постигаем далеко не всё в сознании и мозге, по если бы мы могли устранить наше невежество, решение этой проблемы сразу и со всей очевидностью раскрылось бы пред нами» — McGinn C. *The Making of a Philosopher*. P.181.

⁷³ См., напр.: McGinn C. *The Mysterious Flame: Conscious Minds in a Material World*. N. Y., 1999. P. 72-73. Сходную в некоторых аспектах позицию в 1982 г. занимал Ф. Джексон.

⁷⁴ См.: McGinn C. *The Mysterious Flame*. P. 54-62; McGinn C. *Problems in Philosophy*. P. 37-38, 18-20.

⁷⁵ Внутреннее чувство Макгинн называет интроспекцией. ⁷⁶ См.: McGinn C. *Can we solve the mind-body problem? // The Nature of Consciousness*. P. 530. Первоначально эта статья Макгинна была опубликована в 1989 г. в журнале *Mind*.

⁷⁷ *Ibid.* P. 534. Макгинн подчеркивает, что поддерживает идею Канта о «пространственности формы внешней чувственности», и добавляет, что свойство мозга, знание которого могло бы раскрыть природу его отношения к сознанию, непространственно и «поэтому нуменально относительно чувств*». Далее он отмечает, что понятие реальности должно быть таким, чтобы мыслить «ее подлинно независимой от присущих нам когнитивных способностей» (P. 539; в рамках нашего проекта Международного кантовского интервью Макгинн сказал, что считает главным достижением Канта то, что «тот осознал, что реальность не конституируется ни нашим знанием о ней, ни ее явлением нашей чувственности» — 100 этюдов о Канте. М., 2005. С. 62). Так что свойство, от которого зависит сознание, присуще реальности самой по себе. См. также: McGinn C. *The Problem of Consciousness: Essays Towards a Resolution*. Oxford, 1993 (1991). P. 84.

⁷⁸ Позиция Канта является более осторожной, чем позиция Макгинна, так как, в отличие от Макгинна, Кант не утверждает ничего определенного о специфике отношения сознания и мозга.

⁷⁹ См., напр: Жучков В.А. *Кант и проблема сознания // Философия сознания: история и современность*. М., 2003. С. 52-74.

⁸⁰ См. в этой связи интересные статьи В. Н. Кузнецова — Проблема сознания в философии Жюльена-Офре де Ламетри и «философия сознания» Джона Сёрла // Философия сознания: история и современность. С. 28-35; Проблема «сознание—тело» в философии Нового времени // Философия сознания: классика и современность. М., 2007. С. 18-24.

⁸¹ Далее в этой книге я исхожу из того, что понятие реальности самой по себе, обособленной от чувственных данностей, несостоятельно. С одной стороны, оно вырвано из контекстов обыденного опыта, а с другой стороны, лишено какого-то своего позитивного содержания. Однако это допущение не означает признание эпистемической прозрачности мира. Речь идет лишь о невозможности использования понятия ноуменальной реальности в позитивной аргументации. Впрочем, главные выводы этой книги (см. гл. 5) не зависят от упомянутого допущения.

⁸² См., напр: King P. Why isn't the mind-body problem medieval // *Forming the Mind: Essays on the Internal Senses and the Mind/Body Problem from Avicenna to the Medical Enlightenment*, ed. by H. Lagerlund. Dordrecht, 2007. P. 187-206. Автор доказывает, что возникновение проблемы сознание—тело в современном смысле слова (отсчет ее истории он начинает с Декарта) можно связать с субъективизацией вторичных качеств в науке Нового времени. Оторвавшись от привязки к материи, эти качества, т. е. цвета, запахи, вкусовые, слуховые и тактильные ощущения, стали ментальными, что объективно побуждало спрашивать об их отношении к телам.

⁸³ Или «физического».

⁸⁴ Ключевую роль Декарта иногда оспаривают (см., напр.: Robinson D. N. *Consciousness and Mental Life*. N. Y., 2008. P. 56-57), неоправданно отождествляя историю традиции с историей идей.

⁸⁵ Душа считалась «мыслящей субстанцией», тело — «протяженной субстанцией».

⁸⁶ Приводились и более специальные доводы. Так, Лейбниц утверждал, что Декарт допускал взаимодействие между душой и телом из-за неверного понимания законов сохранения.

⁸⁷ Эта «исправленная» теория естественного влияния, подразумевавшая, что тела, к примеру, не сами производят ментальные состояния, а лишь побуждают душу к этому, восходит к Вольфу, хотя сам он сохранил верность теории предустановленной гармонии. Но ее развил Кнутцен и др.

⁸⁸ См.: Hissmann M. *Psychologische Versuche. Ein Beitrag zur esoterischen Logik*. Frankfurt und Lpz., 1777. S. 248-249; Priestley J. *Disquisitions relating to Matter and Spirit*. L., 1777.

⁸⁹ Главным защитником субстанциального дуализма среди современных аналитических философов (в российской философии таковым оказывается, пожалуй, В.К. Шохин, и о маргинальности дуалистических взглядов в нашем философском сообществе можно говорить лишь условно) является Р. Суинберн, автор «Эволюции души» (1986). Но даже он подчеркивает отличия своей позиции — «мягкого дуализма» — от картезианской версии этой доктрины, отмечая, что считает невозможным высказать что-либо определенное о природе субстанциальной души, — см.: Swinburne R. *The Evolution*

of the Soul. Rev. ed. Oxford, 1997. P. 10. И этот мягкий дуализм сводится к утверждению, что «личности могут — это логически возможно — сохранять свое бытие при прекращении существования тел» — *ibid.* P. 154. Если Суинберн хочет сказать, что из того, что мы можем помыслить ситуацию, когда разрушение мозга не приводит к исчезновению личности, т.е. определенной совокупности сознательных состояний, можно сделать вывод, что сознание существует на ином, нежели мозг, носителе, т. е. является состоянием нематериальной души, то очевидно, что это не вытекает из принятых посылок. Максимум, что из них можно вывести, — что сознание могло бы быть состоянием души. Если же он просто отождествляет душу с совокупностью состояний сознания, то такой дуализм уже настолько «мягок», что не может быть назван субстанциальным, и он неотличим от так называемого дуализма свойств. Другие проблемы субстанциальных дуалистов видны из дискуссии в этой книге: Green J. B., Palmer S.L. (eds) *In Search of the Soul: Four Views of the Mind-Body Problem*. Madison, 2005. P. 61-73.

⁹⁰ Все они — англоязычные философы. И это не случайно. Немецкие и французские философы заметно отстают в этой области. В Германии в последнее время одно за другим публикуются «Введения в философию сознания» (или «духа»), по в них анализируются в основном англоязычные источники. Нечто подобное мы видим и во Франции, только в меньших масштабах.

Глава 2 СЁРЛ: ПРОЩЕ ПРОСТОГО!

Джон Сёрл, как можно подумать, не нуждается в особом представлении — его имя хорошо известно во всем философском мире. Дело в том, что идеи Сёрла так или иначе прозвучали не только в англоязычной аналитической философии, но и в других традициях. Он участвовал в заочной полемике с Жаком Деррида, его философия языка повлияла на Апеля и Хабермаса и т. п. Даже российские авторы, не без труда избавляющиеся от изоляционистских настроений, хорошо знают об этом мыслителе и азартно критикуют его теории¹. Вместе с тем Сёрл остается загадкой для историков мысли. Довольно трудно найти позицию, с которой удалось бы ухватить самую суть его концептуальных построений и, так сказать, прочувствовать пафос его философии. Слово «пафос» я употребил не случайно. Сёрл — очень пафосный мыслитель, автор, не любящий полутонов.

Этот человек просто удивляет своей прямоотой. Например, если ему не нравится какой-то доклад, он может забраться на сцену и попытаться сорвать его². Или сказать все, что он думает о докладчике. Так, свое выступление на конференции в Туссане 7 апреля 2006 г. он начал с поучений в адрес предыдущих ораторов (Д. Розенталя и Р. Ван Гулика), не умеющих, по его мнению, подавать материал. Еще более резким Сёрл бывает в печатной полемике. Деннета, как мы знаем, он уличал в умственной неадекватности³, а влиятельную книгу Д. Чалмерса провозглашал коллекцией абсурдов⁴.

Справедливости ради отметим, что Сёрл беспощаден не только к другим, но и к себе. В частности, он признает свою, скажем так, недостаточную осведомленность в истории дисциплины, которую он 40 лет преподает в университете Беркли (Калифорния, США)⁵. Объясняет он это тем, что медленно читает и что в Оксфорде, где он получал философское образование, не было подробных исторических курсов: «Я изучал историю философии Нового времени, прежде всего Декарта, Локка, Беркли и Юма. Но я вообще не изучал Канта, Лейбница или Спинозу и мало касался Платона и Аристотеля»⁶.

Сёрл, впрочем, пытается истолковать недостаток собственных историко-философских познаний как преимущество, указывая - вполне в картезианском ключе — на то, что своими рассуждениями философы часто не проясняли, а, наоборот, затемняли те или иные вопросы. Созданная ими путани-

ца способствовала тому, что сами по себе несложные проблемы превращались в их глазах в великие тайны.

Именно это, по его мнению, произошло с проблемой отношения ментального и физического. Давняя традиция определенным образом трактовать это отношение, а именно противопоставлять их, была прочно зафиксирована в философской терминологии, задающей рамки современным дискуссиям о сознании. «Почти во всех работах [на эту тему], которые я читал, — замечает Сёрл, — принимается один и тот же набор исторически унаследованных категорий для описания ментальных феноменов, особенно сознания, и вместе с этими категориями — набор допущений о том, как сознание и другие ментальные феномены относятся друг к другу и к остальному миру. И вот этот набор категорий и допущений, которые, словно тяжелый багаж, упомянутые категории тащат с собой, совершенно не подвергается сомнению и способствует продолжению дискуссий. Формирующиеся затем различные позиции все принимаются внутри данного набора ошибочных допущений. Результат таков, что философия сознания занимает уникальное место среди современных философских предметов, и уникальность эта состоит в ложности всех наиболее известных и влиятельных теорий в этой области»⁷.

Неудивительно, что при таком отношении к философской традиции обсуждения проблемы сознание—тело — Сёрл не пользуется чалмерсовским термином «трудная проблема сознания» — он предлагает абстрагироваться от прежних вариантов ее решения и непредвзято взглянуть на фактическую сторону дела: «Мой метод в философии заключается в том, чтобы постараться забыть об истории проблемы, а также традиционных способах мыслить ее, и попытаться ограничиться констатацией фактов, насколько они нам известны»⁸. При таком подходе оказывается, что «знаменитая» проблема, над которой бились лучшие умы, на деле очень проста⁹.

Оценивая позицию Сёрла, стоит обратить внимание на то, что он, с одной стороны, считает, что псевдотрудность проблемы сознание—тело связана с терминологической путаницей, с другой — что она может быть решена при надлежащем учете фактов. Эти на первый взгляд неприметные установки в действительности, как мне кажется, очень показательны, и их анализ позволяет найти ту самую точку зрения — о трудности отыскания которой я говорил в самом начале главы, — с которой можно понять специфику философии Сёрла.

Все дело в том, что упомянутые моменты напрямую связаны с двойственным истоком философских взглядов Сёрла. Как мы уже знаем, их формирование началось в Оксфорде, где он учился в 50-е годы. В те времена этот городок был признанным цен-

тром мировой мысли. Здесь работали такие звезды, как Райл, Остин, Стросон, Айер и др. О Райле Сёрл слышал еще до приезда из США, но его упрощенные лекции, скорее, разочаровали его. Зато Стросон и Остин оказали определяющее влияние на стилистику его философствования.

Философское сообщество долгое время воспринимало Сёрла как прилежного последователя Остина, хотя самому Сёрлу казалось, что наибольшее воздействие на него оказал именно Стросон. Возможно, его ощущение объясняется психологическими причинами: по отношению к Остину он испытывал чувство некоего превосходства — ему представлялось, что он может легко разбить его аргументы, тогда как Стросон в период научного руководства над Сёрлом, напротив, не раз без особых усилий сокрушал выдвигаемые им тезисы¹⁰.

В действительности оба эти влияния имели существенное значение, хотя если смотреть на дальнейшую судьбу Сёрла со стороны, то на поверхностный взгляд более значительным выглядит все же влияние Остина. Ведь Сёрл продолжил ости-новский проект исследования так называемых речевых актов, систематизируя и уточняя идеи учителя. Однако если присмотреться, как именно он уточнял и систематизировал их, то мы увидим, что в общем и целом он делал это скорее в стросоновском, нежели в остиновском ключе.

Так или иначе, но нам не обойтись без нескольких слов о Джоне Остине (1911-1960). Это был настоящий гений. Он мог иногда взглянуть на известные вещи с совершенно новой стороны. Остин не так много публиковался, предпочитая прямое общение со студентами, позволявшее ему погружаться в стихию обыденного языка, который и был основным предметом его изысканий. На семинарах, взяв какое-нибудь философски значимое слово, он выписывал десятки его обыденных употреблений, что позволяло в итоге усматривать его смысловое ядро — на контрасте с близкими словами¹¹. Остина часто сравнивают с поздним Витгенштейном. Сходство и впрямь есть, и может быть, именно поэтому Остин ревниво относился к идолу аналитической философии. Как бы то ни было, подобно Витгенштейну, Остин считал, что решение традиционных философских проблем надо искать в повседневной речи. Т. е. например, если мы хотим понять, можно ли знать о ментальных состояниях другого человека или можно лишь догадываться о них, то ответ содержится в практике употребления глагола «знать». Если выражения типа «я знаю, что ты рассержен» нормальны для обыденной речи, то корректно говорить, что мы знаем о ментальных состояниях других людей.

Только что приведенный пример взят из статьи Остина «Чужое сознание» (1946). Похожие идеи можно найти

и в «Философских исследованиях» Витгенштейна, опубликованных несколькими годами позже, но созданными по большей части задолго до работы Остина. К примеру, в одном из фрагментов Витгенштейн говорит, что проблема «Могут ли машины мыслить?» не стоит раздумий, так как обыденное словоупотребление подсказывает отрицательный ответ¹².

В общем, и поздний Витгенштейн, и ранний Остин были убеждены, что анализ обыденной речи дает ключ к разгадке философских вопросов. Но далеко не все были согласны с таким подходом. В самом деле — если вернуться к примеру со знанием о ментальных состояниях другого человека, — почему мы должны быть уверены, что слово «знать» употребляется в одном и том же смысле, когда мы говорим о собственных состояниях и о состояниях другого человека? Не исключено ведь, что оно употребляется в разных смыслах. Конечно, различие этих смысловых оттенков можно опять-таки эксплицировать в вербальной форме. Но кто сказал, что в обыденной речи содержатся готовые рецепты такой экспликации? Что если об этих реальных различиях никто никогда не думал и не говорил? Тогда, чтобы зафиксировать их, придется изобрести новые слова. Новый, необыденный, технический язык — язык философов — может тем самым способствовать уточнению смысла обыденных выражений¹³.

Это, в свою очередь, означает, что некоторые концептуальные проблемы и средства их разрешения выходят за пределы компетенции обыденного языка и не могут быть распутаны простыми дескрипциями повседневного словоупотребления. А допустить это — все равно что признать, что философия не сводится к разрешению лингвистических головоломок, что она может иметь дело с реальными проблемами. Именно эту мысль пытался провести Карл Поппер во время легендарного спора с Витгенштейном, состоявшегося в один из осенних вечеров 1946 г. в послевоенном Кембридже. Отстаивая свою правоту, Витгенштейн, как известно, был не вполне последователен, пустив в ход раскаленную ко-чергу, которую невозможно причислить к средствам обыденного языка¹⁴. Любопытно, однако, что в пятидесятые годы сюжет этого спора был воспроизведен и с участием Остина¹⁵, оппонентом которого выступил именно Сёрл. На этот раз все было не столь картинно, хотя без эксцессов тоже не обошлось: Сёрл подошел к Остину после его лекции, и они так увлеклись обсуждением вопроса о том, есть ли в философии реальные проблемы, решение которых выходит за пределы возможностей анализа обыденной речи, что не заметили, как служители заперли учебный корпус: им пришлось долго искать ход наружу¹⁶.

Разумеется, я не случайно упоминаю все эти исторические эпизоды. Как мы видим, еще в студенческую пору Сёрл осознавал принципиальные расхождения своих взглядов с воззрениями Остина.

И расхождения эти касались того, что, в отличие от Остина, он полагал, что философия не сводится к анализу обыденной речи. По крайней мере, сам Сёрл представлял и представляет их именно так. Интересно, однако, что в поздней работе «Принесение извинений» (1956-1957) Остин, вполне в сёр-ловском духе, заявлял, что обыденный язык не является для него «последним словом»¹⁷. В принципе, нельзя исключить, что это уточнение было сделано Остином под влиянием упомянутой выше дискуссии с Сёрлом, которая, судя по всему, состоялась в более ранние годы. С другой стороны, трудно не согласиться с А. Строллом, полагающим, что такое расширенное понимание философии предполагается уже остиновской теорией речевых актов¹⁸, о которой еще пойдет речь ниже — которая, впрочем, тоже была разработана Остином в его относительно поздний период¹⁹.

В любом случае Остин четко не обозначил ответ на вопрос, к чему же сводится философия, если она не сводится к анализу речи. Если переадресовать этот вопрос Сёрлу, то мы услышим ответ, и этот ответ состоит в том, что философия сводится к анализу фактов. Разумеется, среди них имеются и факты обыденного словоупотребления, но ими, по Сёрлу, нельзя ограничиваться. Можно сказать, что «факты вообще» занимают у Сёрла место «остиновских» фактов обыденного языка. Однако при всех мнимых и реальных отличиях от Остина Сёрл сохраняет в своих рассуждениях одну типично остиновскую черту. Остин с недоверием относился к истории философии, считая (как и Витгенштейн), что мыслители прошлого скорее создавали проблемы своим произвольным словоупотреблением, чем решали их. Сёрл утверждает нечто подобное: своей искусственной терминологией философы способствовали затемнению фактов. И теперь понятно, откуда проистекает недоверие Сёрла к словарю философии сознания, из какого источника берется его вера, что отказ от старого словоупотребления поможет решить загадку ментального.

Допустим теперь, что мы очистили некую проблему от терминологической шелухи и видим перед собой чистые факты. Как анализировать их, оставаясь в пределах философии и не продуцируя естественнонаучный дискурс? Сёрл разъясняет, что философ должен выявлять «структуру фактов» и что ее выявление нередко совпадает с установлением «необходимых и достаточных условий» этих фактов. Он считает, что это

«классический метод аналитической философии — находить необходимые и достаточные условия»²⁰.

Замечания Сёрла любопытны. Если выявление «структуры фактов» и в самом деле можно было бы назвать классическим методом аналитической философии, так как он может быть атрибутирован Расселу и раннему Витгенштейну, то отыскание необходимых и достаточных условий фактов — более позднее «кантианское» привнесение, связанное с именем Питера Стросона, который вернул аналитическую философию в русло метафизики, которая истолкована им в кантианском ключе²¹ как дисциплина, нацеленная на прояснение концептуальных условий возможности нашего опыта и его предметов.

Итак, если недоверие Сёрла к искусственным терминам идет от Остина, то его метод анализа фактов унаследован им от Стросона. Правда, свойственный работам Стросона кантианский дух чувствуется у Сёрла в гораздо меньшей степени (не будем забывать, что он плохо знает работы Канта). Стоит, впрочем, отметить, что кантовская трансцендентальная методика выявления условий возможности фактов²² - исторически далеко не единственный вариант такого рода анализа. Еще до Канта Христиан Вольф указывал, что задача философии состоит именно в определении условий возможности фактов, причем, в отличие от Канта, Вольф акцентировал внимание на объективных, а не на субъективных условиях возможности²³. И можно сказать, что Сёрл, отдавая должное исследованию субъективных условий, в целом продолжает именно «вольфовскую» линию выявления объективной структуры фактов²⁴. Это обстоятельство позволяет понять истоки интереса Сёрла к онтологической тематике, ведь онтология занимается именно экспозицией структуры реальности. Однако онтология — по самой своей сути системная, т. е. синтетическая дисциплина. Неудивительно поэтому, что аналитические философы зачастую сторонились онтологических изысканий. Сёрл чувствует выпадение ряда своих работ из стандартных аналитических контекстов, но решает проблему самоидентификации с типичной для него прямотой: «Это синтез, осуществленный аналитиком»²⁵.

Возведение концептуального образа мира — завершенного в общем и целом в трактате «Конструирование социальной реальности» (1995) — Сёрл начал еще в 60-е годы, когда он работал над книгой «Речевые акты» (1969). Конечно, поначалу у него не было глобальных планов. Его интересовала философия языка. Сёрл опирался на разрозненные идеи Остина. Остин, как мы помним, неустанно размышлял над особенностями обыденного словоупотребления. Как и поздний Витген-

штейн, он видел с словах некий набор инструментов²⁶, каждый из которых имеет свое назначение и функцию. И на каком-то этапе своих штудий Остин осознал, что с помощью слов можно не только как-то описывать реальность, но и создавать новую реальность²⁷. Знаменитый пример — произнесение будущими супругами определенных слов при бракосочетании²⁸, вроде «я согласен/согласна стать мужем/женой такого-то человека». Озвучивание подобных формул в соответствующей ситуации предназначено, скорее, не для констатации желаний, а для создания реальности брака.

Чтобы подчеркнуть деятельный характер подобных высказываний, Остин назвал их «перформативами». Различие между перформативами и констатациями схватывается уже на интуитивном уровне: ясно, к примеру, что констатации могут быть истинными или ложными, а перформативы — нет (хотя они могут быть удачными и неудачными). Но четко концептуализировать эти различия оказалось непросто. Ведь при желании констатирующие высказывания тоже можно истолковать как особого рода действия. Разница между ними, осознал Остин, скорее, не в том, что одно — действие, другое — нет, а в характере и результатах этих действий, «речевых актов». В любом речевом акте, как он считал, можно выделить три составляющие (или даже говорить о трех актах): (1) локутивную (включающую как фонетический, так и смысловой аспекты), (2) иллокутивную (указывающую на то, что мы, собственно, делаем, произнося ту или иную фразу) и (3) перлокутивную, касающуюся последствий этого акта. К примеру, говоря: «Откройте, пожалуйста, окно!», я (1) произношу осмысленную фразу, (2) прошу собеседника открыть окно, и, в результате, (3) вызываю у него побуждение открыть окно²⁹.

Еще раз подчеркнем, что собственно деятельный аспект речевого акта приходится на второй, иллокутивный момент. Мое действие в только что приведенном примере состояло в просьбе. Разумеется, легко помыслить и другие действия подобного рода. Произнося те или иные фразы, я могу обещать, требовать, заявлять и т. д. Естественно возникает вопрос о принципах классификации иллокутивных актов. Сёрл считает, что Остин не смог отыскать эти принципы (хотя очевидно, что он пытался найти их), так как шел снизу, от конкретных речевых актов, тогда как единственно правильный путь, по Сёрлу, состоит в том, чтобы сначала установить общие условия и базовую структуру речевых актов и лишь затем классифицировать их.

Как же установить базовую структуру речевых актов? Ключ к ее пониманию, по мнению Сёрла, — структура созна-

ния. Ведь всякий речевой акт — это осмысленный звуковой ряд, а осмысленность ему придает именно наложение на него ментальных состояний говорящего. Набор возможных ментальных состояний, которые можно упаковать в звуки или письмо, предопределяет, стало быть, реестр речевых актов. И хотя ментальные состояния разнообразны, это не означает, что их нельзя распределить по неким базовым рубрикам.

Таким образом, правильно понятая философия языка должна быть, по Сёрлу, «ветвью философии сознания»³⁰. И в философии сознания его поначалу интересовали именно вопросы, связанные со структурой и типикой ментальных состояний. Важнейшим свойством ментального он, вслед за Brentano, признавал интенциональность, т. е. направленность на объект, или содержательность. Сущностная корреляционность интенциональных состояний позволила ему выделить два «направления соответствия» между подобными состояниями и соответствующими им (или репрезентируемыми ими, как любит выражаться Сёрл) объектами: от сознания к миру и от мира к сознанию. В первом случае «условием сатисфакции» ментальных состояний является соответствие их предметного контента реальному положению вещей, во втором — наоборот. Иначе говоря, ментальные состояния либо отражают мир, как, скажем, убеждения, либо, подобно желаниям, нацелены на вмешательство в него.

Как нетрудно догадаться, Сёрл утверждает, что это базовое различие в «направлении соответствия» интенциональных состояний и мира проявляется в речевых актах. Убеждения и другие интенциональные состояния первого типа выражаются в «ассертивах», а желания и подобные им состояния второго типа — в («директивах» (приказах) и «комиссивах» (обещаниях). Помимо этих типов речевых актов, по Сёрлу, могут существовать еще только два: с вырожденным, или нулевым, направлением соответствия и с двойным направлением. И они действительно существуют, это «экспрессивы» (к примеру, «Извините за оговорку») и «декларативы» (к которым Сёрл относит, в частности, и остиновские «перформативы»). В первом случае мы видим нулевое отношение соответствия, ибо здесь «уже предполагается истинность пропозиционального контента»³¹; во втором — имеется «двойная направленность соответствия, так как мы изменяем мир и таким образом достигаем направления соответствия от мира к слову, путем представления его изменившимся, а значит, достижения направления соответствия от слова к миру»³².

В общем, сложно не согласиться с Сёрлом: философия сознания в состоянии фундаментализировать философию языка, давая ей путеводную нить для классификации речевых актов. Заме-

тим лишь, что сам Сёрл не сразу осознал эту основополагающую роль философии сознания. В «Выражении и значении» и тем более в «Речевых актах» он только нащупывал эффективные подходы к упомянутой классификации. Приоритеты были окончательно расставлены им лишь в книге «Интенциональность», опубликованной в 1983 г. Во введении к ней Сёрл заявил, что рассматривает это «Эссе по философии сознания» (как гласит подзаголовок этой работы) в качестве фундамента для «Речевых актов» и «Выражения и значения» (1979). Но в «Интенциональности» Сёрл почти не выходит за пределы того, что сам он называет логическим анализом интенционального отношения. В этой книге, по сути, не детализируется «феноменология» интенциональности, т. е. ее сознательный, субъективный аспект. Отмечу, что при этом Сёрл все же указывает на связь интенциональности и сознания. Кроме того, в последней главе с характерным названием «Эпилог: интенциональность и мозг» он ставит (и быстро снимает) вопрос об онтологическом статусе сознания. Позже, начиная с программной работы «Новое открытие сознания» (1992), Сёрл все больше рассуждает именно о субъективной феноменологии интенциональных состояний (перечисляя, в частности, такие феноменальные свойства сознания, как «единство», «качественность»³³, «настроение», «гештальтная структура», «чувство Я» и т. п.) и, соответственно, фокусирует внимание на проблемах онтологии сознания.

Таким образом, стартовав с пограничных областей философии сознания, Сёрл постепенно смещался к ее центральной проблеме, касающейся уточнения онтологического статуса ментального. Это движение соответствовало его интересу к выявлению «структуры фактов» и позволило выстроить цельную онтологическую схему. На вершине пирамиды сущего находятся социальные факты. В своей развитой форме они предполагают речевые акты, приписывающие определенным физическим сущностям некий дополнительный статус и соответствующие ему функции (например, прямоугольным бумажкам — статус денежных купюр). Условием речевых актов являются интенциональные структуры сознания, причем их непосредственным конституэнтom оказываются структурные, или «логические», параметры интенциональных состояний, вроде направления соответствия. Однако сама логическая матрица интенциональности предполагает существование субъективной, психической реальности — вне последней, по Сёрлу, можно рассуждать только о «как бы интенциональности», или же о производной интенциональности.

Наконец, условием существования феноменальной сферы сознания, или субъективной реальности, являются, по Сёрлу,

определенного рода физические сущности, биологические объекты, наделенные нервной системой. Это, однако, не должно затушевывать единство биологических систем с остальной материей, состоящей из одних и тех же компонентов. Данный уровень — элементарных частиц — Сёрл считает предельной основой бытия. И его онтологический проект призван показать, как на этой основе, не нарушая физического единства мира, выстраиваются факты более высокого типа, от элементарных ментальных состояний до хитроумных социальных конструкций, таких, как деньги, партии или президенты.

Теперь, оценив широту замысла Сёрла, мы можем присмотреться к ключевому звену его онтологической схемы, а именно к тому решению проблемы сознание—тело, которое он избирает. Мы знаем, что Сёрл предлагает решать эту и подобные ей проблемы, обнажая от терминологической скорлупы релевантные факты, которые при таком подходе сами подсказывают свою внутреннюю структуру.

Начинается самое интересное. Какие же факты, по Сёрлу, следует учитывать при уточнении онтологического статуса сознания? Более или менее определенный ответ на этот вопрос можно найти в его статье «Биологический натурализм», опубликованной в «Блэквелловском справочнике по сознанию» (2007). Но поскольку это очень краткий текст, обсуждение позиции Сёрла лучше проводить на материале его книг, хотя и сверяясь с упомянутой статьей. И в качестве стартовой площадки уместно выбрать его систематический труд «Сознание, язык и общество», впервые опубликованный в 1998 г.³⁴ Сёрл многообещающе начинает обсуждение нашей проблемы: «Всегда неплохо напомнить себе о фактах, напомнить себе о том, о чем мы действительно знаем». Затем он, казалось бы, переходит к их перечислению: «Нам фактически известно, что все наши сознательные состояния каузально порождены процессами в мозге»³⁵. Но дальше все размывается, и приходится догадываться, что вторым фактом он предлагает признать то, что сознание как продукт указанных процессов — это особый феномен. Его отличительная черта — субъективность, означающая, что его можно включать в онтологические построения только «от первого лица», но не «от третьего»³⁶.

Собственно, это и есть основные факты. Теперь надо выявить их соотношение. Напомним, что, по Сёрлу, это совсем несложно, надо только абстрагироваться от традиции жесткого противопоставления ментального и физического. В свое время, в эпоху Декарта, это противопоставление было оправданно, так как с его помощью тот же Декарт уводил физическое из сферы влияния теологии, откупаясь от нее признанием существования нематериальной души. Но дело не только

в Декарте. Психофизический дуализм, считает Сёрл, входит в каркас европейской культуры, в число ее «установок по умолчанию».

О его понимании установок по умолчанию³⁷ стоит сказать подробнее. Уже отмечалось, что большинство ментальных состояний, по Сёрлу, интенциональны. Но интенциональные состояния не существуют в изоляции друг от друга. Они связаны в некую Сеть. К примеру, желание чего-то предполагает множество убеждений в том-то и том-то. Но далеко не все убеждения существуют в эксплицированной форме. Подобные неэксплицированные убеждения, образующие Фон наших интенциональных состояний, уже нельзя называть подлинными интенциональными состояниями³⁸. Некоторые из них точнее охарактеризовать как умения (врожденные или приобретенные), другие же являют собой фундаментальные допущения³⁹, задающие контекст для самых разнообразных теоретических позиций.

Доинтенциональные состояния второго рода Сёрл и называет установками по умолчанию⁴⁰. Любопытно, что он не дает развернутой классификации этих установок и избегает четких высказываний относительно их происхождения. Можно лишь предположить, что некоторые из них — вроде допущения внешнего мира, независимого от сознания, — он считает продуктом биологической эволюции. Вместе с тем он дает понять, что большинство подобных установок имеет не биологическое, а культурное происхождение. Такие установки, кстати, могут сталкиваться между собой. Собственно, одной из функций философии, по Сёрлу, должно быть разрешение конфликта между различными установками по умолчанию, отыскание среднего пути.

Сёрл убежден, что проблема сознание—тело являет собой яркий пример столкновения установок по умолчанию⁴¹. С одной стороны, люди склонны считать сознание чем-то противоположным телу, это дуалистическая установка, восходящая к истокам европейской цивилизации и обостренная Декартом. С другой — они полагают, что мир исчерпывается материальными формами бытия. Усиление этой установки тесно связано с прогрессом естествознания в Новое время. Обе эти установки в своем безусловном выражении ошибочны. Дуализм отрицает единство мира, все более полно раскрываемое наукой, а материализм игнорирует факт существования сознания как особой данности, не описываемой в материальных терминах. Но поскольку и материалистическая, и дуалистическая установки базируются, по Сёрлу, не на глубинных биологических диспозициях, а вырастают на культурной почве (поэтому в некоторых культурных традициях, утверждает он,

проблема сознание—тело даже не может быть сформулирована⁴²), то они, так сказать, безусловны и могут быть отброшены при отыскании лучшей альтернативы. А еще лучше сразу отбросить их — для отыскания таковой.

К проблеме онтологических установок, формирующих наше понимание отношения ментального и физического, мы еще вернемся — совершенно иначе истолковав их природу и действие⁴³. Пока же последуем за Сёрлом и попробуем наконец понять его взгляды на то, как же все-таки следует мыслить отношение сознания и мозга, счастливо избегая Сцилл и Харибд материализма и дуализма.

Итак, Сёрл полагает, что, очистив наш ум от устаревшего картезианского словаря, противопоставлявшего сознание и мозг, мы увидим, что к фактам (1) каузального порождения мозгом сознания как (2) особой субъективной данности (т. е. такой данности, наличие которой предполагает субъекта), можно добавить и еще один, а именно что (3) ментальные, сознательные состояния есть особые свойства мозга, биологические феномены, во многом сходные, к примеру, с секрецией или фотосинтезом. Таким образом, «все формы сознания каузально порождены поведением нейронов и реализованы в системе мозга, которая сама состоит из нейронов»⁴⁴. Это, провозглашает Сёрл, и есть искомый путь между материализмом и дуализмом. Он именуется своей позицией «биологическим натурализмом».

Вот, собственно, и все. «Неужели решение знаменитой "проблемы сознание—тело" действительно может быть столь простым?» — восклицает Сёрл, предвидя раздражение читателей⁴⁵. Может, уверен он, если непредвзято смотреть на нее. С одной стороны, исключить, что он прав, конечно, нельзя. С другой — иногда говорят, что у любой трудной проблемы есть простое решение, и притом всегда неверное. Нам еще предстоит увидеть, применимо ли это к Сёрлу, но сейчас я хочу обратить внимание на другое: по правде говоря, на первый взгляд не очень ясно, в чем состоит его решение!

Эту неясность, впрочем, можно устранить. Говоря о порождении мозгом субъективных ментальных состояний, которые, как подчеркивает Сёрл, онтологически нeredуцируе-мы к объективным физическим свойствам нейронных систем, он избегает опасности материализма, ведущего к отрицанию сознания и конфликтующего со здравым смыслом. Утверждая же, что сознательные состояния реализованы в мозге, он отмежевывается от дуализма, согласно которому ментальное является самостоятельным регионом бытия.

Надо уточнить, что Сёрл атакует сразу две разновидности дуализма: субстанциальный дуализм и дуализм свойств. Суб-

станциальный дуализм исходит из того, что ментальные состояния — свойства так называемой «духовной субстанции», а не мозга. Этот вид дуализма Сёрл мог бы исключить сразу - ведь он считает фактом каузальное порождение мозгом сознания. Впрочем, отдельные философы — прежде всего я имею в виду У. Хескера⁴⁶ — в наши дни считают, что субстанциальный дуализм можно совместить с тезисом о порождающей роли мозга, т. е. они утверждают, что мозг порождает духовную субстанцию. Любопытно, кстати, что эта теория не так уж враждебна идеям Сёрла. Напомню, что он определяет важнейшее свойство субъективности сознательных состояний через отсылку к субъекту: «...сознательные состояния субъективны в том смысле, что они всегда переживаются человеческим или животным субъектом»⁴⁷. Но что понимается здесь под субъектом? Это едва ли мозг, так как от существования мозга зависит и множество сугубо физических свойств этой системы, но Сёрл не считает их субъективными. Не может он и сослаться на то, что эти физические свойства, в отличие от сознательных состояний, не «переживаются» мозгом. Ведь тогда сознание будет определяться через «переживание», но это ничего не решит в вопросе о его субъективности, если только мы не скажем, что переживание обеспечивает доступ к приватным данностям, — но тогда понятие субъекта станет лишним элементом объяснительной схемы⁴⁸. Долгое время Сёрл, по сути, не прояснял природу субъекта, но в работах 2000-х гг. он старался восполнить этот пробел, и он близок к выводу, что субъект, или самость (Я), есть некая инстанция, отличная как от мозга⁴⁹, так и от зависимых от него сознательных состояний, хотя, по-видимому, и порожденная мозгом, переживающая сознательные состояния, поддерживающая единое поле сознания и служащая источником свободных решений⁵⁰.

Вместе с тем, теории Хескера и Сёрла на деле мало отличаются от дуализма свойств. Дуализм свойств тоже признает каузальную зависимость ментальных состояний (которые могут быть и тесно интегрированными) от мозга, но его сторонники убеждены, что ментальные состояния, будучи особой реальностью, не могут быть онтологически редуцированы к физическим аспектам мозга, т. е. не могут отождествляться с ними. Поскольку Сёрл считает так же (оговариваясь при этом, что под «физическим» он понимает пространственные характеристики мозга, изучаемые нейронаукой), некоторые авторы, к примеру Д. Чалмерс и Дж. Ким, полагают, что он фактически примыкает к этой доктрине⁵¹. Сёрл, однако, не согласен с таким выводом⁵². Но как избежать его? Он опробует интересное решение, раскрываемое как раз через поня-

тие «реализованности в мозге» ментальных состояний. Она означает — и это главный «структурирующий» ход Сёрла в трактовке сознания, — что ментальные состояния могут быть поняты как высокоуровневые, «эмерджентные» свойства нейронных систем⁵³. Иными словами, сознание — это не особая ментальная реальность, а специфический уровень физической реальности.

Если это так, если сознание действительно может быть поставлено в один ряд с такими эмерджентными свойствами (т. е. свойствами, не присущими компонентам некоей системы, но присущими системе в целом), как, например, текучесть воды, то мы, по Сёрлу, получаем возможность уяснить еще один аспект проблемы сознание—тело, связанный с возможностью ментальной каузальности. Суть в том, что еще одним фактом, по Сёрлу, является то, что наше сознание оказывает влияние на наше поведение⁵⁴. Скажем, если мы желаем чего-то, то это желание толкает нас на определенные поступки. Но если ментальные состояния обособлены от физических процессов в нашем теле, то это влияние мыслимо лишь при нарушении физических законов. Признать подобное влияние, считает Сёрл, значит идти против науки: совершенно бессмысленное предприятие. Другое дело, если ментальные состояния — высокоуровневые свойства физических систем. Тогда они могут иметь реальную каузальную действенность, хотя эта действенность не выходит за пределы каузальной действенности низкоуровневых компонентов соответствующей системы, в данном случае — нейронов. Иными словами, желание, скажем, попробовать медовуху — это высокоуровневое свойство определенных нейронных процессов в мозге. Причины поведения, вызываемого этим желанием, могут быть описаны исключительно в физиологических и биохимических терминах, так же как «поведение» воды может быть описано в молекулярных терминах. Но эти причины можно описать и на более высоком уровне - уровне желаний, в случае мозга, и уровне текучести — в случае воды.

Теперь мы видим, что подход Сёрла к проблеме сознание-тело и в самом деле не лишен достоинств. Главное из них - удивительное согласие его решения со здравым смыслом. Если попросить не ангажированного оккультизмом (т. е. не утратившего здравый смысл) «человека с улицы» порассуждать о сознании, то мы услышим, что сознание — это какое-то внутреннее чувство, что сознание зависит от мозга и что оно влияет — и еще как — на поведение. Концептуальная схема Сёрла объединяет все эти здравые тезисы.

В общем, Сёрл — рыцарь здравого смысла в философии сознания⁵⁵. Однако, как учит история, это не только в чем-то

даже беспроигрышная, но и довольно рискованная роль. В свое время Томас Рид основал целую школу «здорового смысла», чтобы разрушить утонченные теории Дэвида Юма. То противостояние вызвало много откликов. За Юма вступился даже Кант, обвинивший его противников в недооценке его идей. Кант был убежден, что Юм ставил реальные проблемы, которые надо было решать, а не отмахиваться от них.

Думается, что Кант тонко почувствовал тенденцию философов здорового смысла скользить, так сказать, по поверхности понятий, не вдаваясь в серьезный концептуальный анализ. Правда Канта становится еще более очевидной, если принять во внимание одну деталь, проливающую свет на соотношение позиций Юма и Рида: подлинной предметной областью философии, по Юму, является именно обыденная жизнь, суждения здравого рассудка, и философия должна прояснять и, где это возможно или необходимо, подправлять их. Таким образом, Юм, ничуть не меньше Рида, предлагал заниматься принципами здравого смысла. Но если Рид считал их последней инстанцией, к которой можно апеллировать, то Юм, соглашаясь с тем, что они играют решающую роль в повседневной жизни, требовал подвергнуть их исследованию, и вовсе не затем, чтобы отбросить их — это все равно в конечном счете невозможно, — а просто для того, чтобы понять, как функционирует обыденный рассудок, откуда возникают принципы здравого смысла и как они соотносятся между собой.

Размышляя о философии Сёрла, мы замечаем те же тревожные черты, которые были характерны для Рида и его последователей. Перечисляемые им «факты», имеющие отношение к проблеме сознание—тело, являются просто расхожими суждениями на этот счет. И хотя все они, возможно, истинны — и, как увидят внимательные читатели в конце книги, я и правда считаю многие из них таковыми, — но это не значит, что их не надо подвергать тщательному анализу. При его отсутствии мы просто не в состоянии понять соотношение этих фактов, и все утверждения на этот счет останутся декларациями. Вместо реального понимания их соотношения мы получим слова, скрывающие, но никак не устраняющие концептуальные провалы.

Демонстрация этих провалов не потребует больших усилий. Начать хотя бы с первого сёрловского факта о каузальном порождении мозгом сознания. В книге «Тайна сознания» (1997) Сёрл бросает: «Если мы вообще что-то знаем о мире, то мы действительно знаем, что процессы в мозге каузально порождают наши сознательные состояния»⁵⁶. Однако буквально на предыдущей странице той же книги он рассуждает о том, как трудно определить конкретные корреляты и причины тех

или иных ментальных состояний и как мы еще далеки до достижения этой цели (позже я постараюсь показать, что она, в том виде, как Сёрл понимает ее, вообще недостижима).

Как же может быть так, что нам совершенно неизвестны корреляты или причины конкретных ментальных состояний, но при этом мы совершенно уверены, что мозг каузально порождает их? Что стоит за этой уверенностью? Очевидно, что за ней стоит тот простой факт — и это действительно факт, — что изменения, происходящие в мозге, как правило, сопровождаются сменой ментальных состояний. Но не требуется труда, чтобы показать, что этот факт совершенно недостаточен для утверждений о каузальной роли мозга в порождении сознания. Ведь он может быть согласован и с другими гипотезами, к примеру с гипотезой предустановленной гармонии, подразумевающей дуалистическую схему соотношения сознания и тела и простой параллелизм физических и ментальных состояний. Разумеется, никто больше всерьез не принимает эту гипотезу, но речь идет о логической непротиворечивости дуализма. Если дуалистическая система не содержит логических противоречий, факт соответствия ментальных состояний и состояний мозга не может быть равносителен признанию каузальной зависимости первых от вторых.

Сёрл, кстати говоря, признает, что субстанциальный дуализм «логически возможен»⁵⁷, и это означает, что его тезис о факте порождения сознания мозгом некогерентен даже внутри его системы, если понимать под фактом данность, констатация которой не вызывает сомнений. Впрочем, я не хочу сказать, что это возражение разрушает построения Сёрла. В конце концов, я только что отметил, что альтернативное объяснение соответствия ментальных и физических состояний - через гипотезу предустановленной гармонии - в наши дни уже никем не принимается всерьез. Но почему? Не из-за его логических противоречий, а просто потому, что это объяснение требует расточительного нагромождения недостоверных сущностей — души и Бога как верховного координирующего принципа. Гораздо более последовательно мыслить ментальное как продукт физического.

Таким образом, упрек Сёрлу по поводу его базового « факта» о порождении мозгом сознания сводится, скорее, к тому, что он надлежащим образом не исследует это положение и не выявляет все его компоненты. Такое выявление может показаться излишеством, но лишь на первый взгляд. Детали могут иметь решающее значение. И прояснить нужно в том числе понятие причины, чтобы быть уверенным в том, что именно это понятие лучше всего подходит для характеристики отношения мозга и сознания. Сёрл немало говорит о причинности, но его рас-

суждения могут скорее запутать ситуацию, чем прояснить ее. При первом приближении может показаться, что он разделяет юмовское толкование причины как постоянного коррелята некоего события, предшествующего во времени последнему. В действительности Сёрл не считает предшествование во времени существенным компонентом понятия причины и утверждает, что это понятие может работать и для одновременных феноменов. Он считает, к примеру, возможным говорить о том, что молекулы воды являются совместной причиной ее текучести, и точно так же допустимо говорить о сознании, являющемся действием некоего множества нейронов.

При такой трактовке понятие действия как события, вызванного каким-то другим событием, причиной, радикально трансформируется и становится синонимичным понятию свойства. Вместе с тем Сёрл продолжает говорить о корреляции как одной из характеристик причины. Обратим внимание, как плохо стыкуется это понятие с понятием свойства: язык сильнейшим образом противится рассуждениям о множестве молекул воды как корреляте текучести. И понятно почему. Свойства, подобные текучести, — это просто предрасположенности неких групп молекул вести себя определенным образом. Эта предрасположенность логически вытекает из структуры последних⁵⁸, подобно тому как из структуры одинаковых равнобедренных прямоугольных треугольников логически вытекает то, что они могут образовывать квадрат. Эмерджентное свойство квадратообразования (это действительно эмерджентное свойство, так как оно не присуще ни одному из отдельно взятых треугольников) нельзя назвать коррелятом двух таких треугольников. Ведь коррелят — это феномен, реально отличный от другого феномена и сопровождающий его.

Иными словами, коррелят, в отличие от следствия, нельзя логически вывести из того, коррелятом чего он является. И теперь понятно, почему Сёрл проявляет столь драматичные колебания в характеристике сознания и его нейронной основы. Трактую сознание по аналогии с текучестью, он должен был бы склоняться к тезису о логической выводимости сознания из его нейронной основы (т. е. к тезису, что при определенном сочетании нейронных процессов невозможно отчетливо помыслить отсутствие сознания). Однако вместе с тем он признает, что феномен сознания онтологически нетождествен его нейронной основе и, соответственно, логически не выводим из нее. Поэтому он сохраняет его характеристику как коррелята.

Но, утверждая коррелятивность сознания нейронным процессам, он — чтобы сохранить элементарную последователь-

ность — больше не должен был бы рассматривать ментальные состояния в качестве аналогов эмерджентных свойств материи, таких, как текучесть. Однако отказ от этой аналогии нанес бы сокрушительный удар по его объяснению ментальной каузальности. Действительно, если текучесть как свойство воды есть некий «логический результат»⁵⁹ более фундаментальных параметров ее частиц, то можно утверждать, что это свойство ответственно за какие-то действия, хотя в принципе причины последних можно описать и на более низком уровне — так же, как «порождение» десятки можно описать на «высоком» уровне как результат сложения двух пятерок или же на "низком", при котором каждая пятерка будет рассматриваться как сумма пяти единиц. «Каузальная» сила пятерки тождественна каузальной силе пяти единиц. Но все эти схемы не работают, когда речь идет о сознании. Одно дело выдвигать альтернативные описания одних и тех же процессов, и совсем другое — говорить о разных процессах. Ведь Сёрл признает онтологическую обособленность ментальных состояний от нейронных процессов в мозге. Ментальные состояния — одно, нейронные процессы — нечто другое. Поэтому он не может утверждать, что, к примеру, желание — это нейронный процесс на высоком уровне описания. Ведь описание в широком смысле - это картина какого-то феномена. Но желание — это не картина, а феномен.

Все это означает, что Сёрл должен был бы сделать принципиальный выбор. Если он держится за тот «факт», что сознание играет реальную каузальную роль в детерминации человеческого поведения, то он должен был либо сказать, что ментальные состояния тождественны тем или иным состояниям мозга (и тогда, разумеется, они могут играть реальную каузальную роль), либо, допуская их различие, утверждать, что их каузальная действительность выходит за пределы каузальной действительности нейронных процессов⁶⁰. Первый вариант противоречит тезису об онтологической нередуцируемости качественных состояний, второй — если принять, что физическая реальность дается с позиции третьего лица, — приводит, как кажется, к выводу о нарушении физических законов и принципа каузальной замкнутости физического⁶¹ вследствие активности нашего сознания.

В действительности с последним пунктом все далеко не так однозначно, и мы еще вернемся к этому вопросу в конце книги. Сейчас достаточно отметить, что сам Сёрл считает такое нарушение неизбежным и не хочет занимать такую позицию, которая его допускает⁶². Неприемлем, повторим, для него и тот вариант, который подразумевает отождествление ментальных состояний и нейронных процессов. Соответствен-

но, он должен был бы отказаться от тезиса о том, что сознание оказывает реальное влияние на поведение. Но отказаться от него — значит, считает он, пойти на противоречие «факту», точнее положению здравого смысла о наличии такого влияния. Я не случайно ставил слово «факт» в кавычки. Разумеется, это никакой не факт в строгом смысле слова. Здравый рассудок и правда говорят нам, что ментальные состояния в какой-то степени ответственны за наше поведение, но эта уверенность может быть простой иллюзией, возникающей потому, что этому рассудку ничего не известно о нейронных процессах, определяющих действия человека.

Что же решает Сёрл? Какой выбор он делает? Очень нестандартный, надо признать. Он не отказывается ни от чего и провозглашает: (1) сознание влияет на поведение, будучи (2) высокоуровневым свойством нейронных процессов (это логично), но таким свойством, которое (3) онтологически отлично от этих процессов⁶³.

Проблема в том, что третий тезис так конкретизирует второй, что понятие высокоуровневого свойства в нем больше не может рассматриваться как альтернативное описание нейронных процессов, которые, как предполагается, определяют поведение. Отсюда следует, что это свойство должно признаваться обособленным феноменом, не играющим каузальной роли, т. е., иными словами, эпифеноменом. Это, в свою очередь, приводит к противоречию с первым тезисом. А это значит, что Сёрл лишь на словах решает проблему ментальной каузальности, на деле же его позиция попросту некогерентна.

Некогерентность позиции Сёрла⁶⁴ — неприятный сюрприз для сторонников простых решений, армию которых он возглавляет. Но нельзя ли все-таки найти какой-то выход из ситуации? Один вариант есть. Сёрлу надо было бы признать, что ментальные состояния и нейронные процессы онтологически тождественны, но эпистемологически различны, вследствие того, что люди обладают двойным доступом к этим нейронным процессам — извне и изнутри. Поэтому возникает ощущение двойственной реальности, на деле этой двойственности нет. Некоторые тексты Сёрла наводят на мысль, что он обдумывал этот, как может показаться, спасительный для него путь. В одной своей поздней книге он замечает, что «сознание есть аспект мозга, аспект, состоящий в онтологически субъективных переживаниях»⁶⁵. «Аспект» — показательный термин, отсылающий к таким эпистемологическим понятиям, как перспектива или взгляд. Но и эту линию Сёрл четко не проводит — хотя, возможно, все еще впереди.

Впрочем, до него эту линию уже провели другие (так что сейчас нас ожидает небольшая историческая экскурсия, ко-

торая, я вынужден признать, несколько сместит основную нить главы). Их было так много, что трудно зафиксировать ее начало. Ясно лишь, что подготовительная работа была проведена Э. Махом и Р. Авенариусом, утверждавшими, что одни и те же элементы бытия предстают в физическом или психическом облике в зависимости от аспекта их рассмотрения. Эти рассуждения были продолжены У. Джемсом и Б. Расселом, проponentами так называемого «нейтрального монизма».

Все упомянутые авторы полагали, что нейтральные элементы могут быть непосредственно даны в опыте. Скажем, дано нечто зеленое. Само по себе оно нейтрально, но если рассматривать его по отношению к нервной системе, то зеленое оказывается ощущением. Физическим же объектом оно оказывается при соотнесении с элементами, не принадлежащими множеству элементов нервной системы.

Такая схема более или менее работает с чувственными данными, но она дает сбои при анализе, скажем, эмоций или образов. Образ, к примеру, как нечто субъективное, приватное, может быть истолкован в физикалистском ключе лишь при отсылке к его гипотетическим нейронным коррелятам. Но невозможно зафиксировать в опыте такую сущность, которая оказывается образом при одной трактовке, нейронным процессом в мозге — при другой. И если мы хотим сохранить монистический, «нейтралистский» подход даже в таких случаях, мы должны интерпретировать эту нейтральную сущность в кантианском смысле, как нечто выходящее за пределы опыта.

К подобному решению склоняется, в частности, К. Мак-гинн. Трудности, возникающие при такой интерпретации, хорошо известны и связаны прежде всего с проблематичностью самого понятия нейтральной реальности, выходящей за пределы возможного опыта. Неудивительно поэтому, что в философской литературе уже давно был опробован совершенно иной вариант решения проблемы возможного тождества ментальных и физических состояний. Он состоял в намеренном отказе от попыток конкретизировать детали искомого тождества. Вместо этого сторонники «теории тождества» акцентировали внимание на аргументах, которые должны приводить к выводу о наличии тождества ментальных и нейро-нальных состояний.

Знаменитая «теория тождества» — концептуальное творение 50-60-х гг. XX века. У. Плейс, Дж. С্মарт, Д. Льюис, Д. Армстронг, Б. Мэдлин и другие соавторы этой теории испытали влияние британского философа Гилберта Райла (1900-1976). Начав как сторонник идей Венского кружка и раннего Витгенштейна и продолжив их попытки выявлять

в философских текстах категориальные бессмыслицы, на зрелом этапе он сместился к лингвистической философии, заставляющей вспомнить о работах позднего Витгенштейна. Кстати, Райл дружески общался с Витгенштейном в 30-е и 40-е годы. Это общение, среди прочего, привело к тому, что он перенес в ретроградный Оксфорд⁶⁶ новые кембриджские методы и их развитие вскоре превратило Оксфорд в столицу аналитической философии.

Нельзя, впрочем, говорить, что Райл во всем следовал позднему Витгенштейну. Оба они анализировали обыденный язык и через этот анализ отсекали бессмысленные философские конструкции. Но если поздний Витгенштейн считал, что бессмыслица возникает тогда, когда слова покидают обыденные контексты, и не связывал их обесмысливание с нарушениями «логического синтаксиса» (такой подход был характерен для него, скорее, в 20-е годы и раньше), то Райл в чем-то остался верен прежней аналитической философии и диагностировал «категориальные ошибки», скрывающиеся за невниманием философов к обыденному словоупотреблению. Райл долго размышлял о том, на каком примере продемонстрировать подобные категориальные ошибки. Сначала, как известно, он хотел взяться за вопрос о свободе воли, но затем остановился на проблеме сознания. Так и появилась книга «Понятие сознания» (1949), которой суждено было стать не только образцом аналитической философии оксфордского периода, но и знаменем философского бихевиоризма, интерес к которому Райл тоже мог унаследовать от Витгенштейна.

Как и полагается радикальному бихевиористу, Райл попытался избавиться от понятия самодостаточных частных ментальных состояний, переопределяя их в терминах поведенческих диспозиций. Скажем, желание выпить воды можно интерпретировать как предрасположенность человека к совершению каких-то действий — к тому, чтобы взять бокал с водой, поднести его ко рту и т. п. И такую предрасположенность совсем не обязательно понимать как какое-то субъективное состояние, это поведенческий кондиционал, периодически возникающий и исчезающий. Избавляя человеческий организм от «привидения в машине», как он называл декартовскую *res cogitans*, т. е. всю частную сферу сознания, Райл, однако, сталкивался с тем, что некоторые внутренние состояния, вроде ощущений или образов, плохо переводились на язык диспозиций. Представление о красной корове, к примеру, это диспозиция к чему? В общем, у поклонников Райла — а их было множество — возникло чувство, что его проект изгнания сознания из тела или, точнее, слияния сознания с поведением, остался незавершенным. И через несколько лет

после «Понятия сознания» у молодого британского психолога и философа Уллина Плейса (в прошлом ученика Райла) возникла идея, как довершить дело учителя. Итогом его размышлений стала статья «Является ли сознание процессом в мозге?», опубликованная в 1956 г. Она-то и положила начало влиятельной традиции теории тождества.

Решение, найденное Плейсом, заключалось в тезисе, что физикализация тех ментальных состояний, которые с трудом поддавались интерпретации в качестве поведенческих диспозиций, может быть осуществлена иначе - путем прямого отождествления этих состояний с состояниями мозга⁶⁷. При этом Плейс не отрицал самого существования субъективных состояний. Он просто говорил, что они могут быть идентичны каким-то процессам в мозге. Он считал это утверждение «научной гипотезой», которая может быть подтверждена при отыскании нейронных коррелятов таких состояний. Искушенные читатели сразу заметят, что критерий верификации, предложенный Плейсом, явно недостаточен, так как отыскание нейронных коррелятов, к примеру, представления о красной корове, никак не меньше может подтверждать и такие взгляды на ментальное, согласно которым оно сопутствует физическому, но не совпадает с ним.

Недоработки Плейса решил устранить Джек Сمارт, который опирался также на идеи выходца из Венского кружка Г. Фейгла, опубликовавшего в 1958 г. эссе «Ментальное и физическое»⁶⁸. Статья Смарта «Ощущения и процессы в мозге» появилась в 1959 г. Он оспорил тезис Плейса о том, что теория тождества имеет характер научной гипотезы, заявив, что «никакой мыслимый эксперимент не может решить дело в пользу материализма или эпифеноменализма»⁶⁹, которые, по его мнению, должны рассматриваться в качестве альтернатив, имеющих при корреляции между сознательными состояниями и процессами в мозге. Выбирая в пользу материализма, т. е. тождества субъективного «опыта» с процессами в мозге, мы следуем оккамовскому принципу простоты. Ведь эпифе-номеналистская теория нуждается в неуклюжем довеске: законах, сопрягающих физические процессы и отдельно от них существующие — в роли «помологических бездельников»⁷⁰ — ментальные состояния.

Важно отметить, что в статьях Плейса и особенно Смарта была намечена и принципиально иная стратегия отождествления ментального и физического, которую можно интерпретировать в качестве продолжения элиминативистской райловской линии⁷¹ — позже мы увидим, что теоретики тождества были, по сути, обречены на то, чтобы проводить обе линии рассуждения, — но не в силу их дополняющих друг дру-

га достоинств, а из-за недостатков обеих. Так вот, Смарт обратил внимание, что предположение о тождестве ментального и физического можно, как кажется, опровергнуть следующей логикой: возьмем какой-нибудь ментальный образ, к примеру образ желтоватого апельсина, и спросим себя: если он тождествен какому-то процессу в мозге, то не будет ли это означать, что этот процесс тоже должен быть желтоватым? Это абсурдно, но как выйти из тупика? Смарт предлагает решение. Во-первых, говорит он, процессу в мозге тождествен не сам ментальный образ, а «опыт его наличия». Во-вторых, опыт наличия образа можно описать на «нейтральном языке»: его наличие означает, что «происходит нечто подобное тому, что происходит, когда мои глаза раскрыты, я не сплю и передо мной находится хорошо освещенный апельсин»⁷². В этом описании ничего не говорится о природе происходящего — и соответственно ничто в принципе не мешает отождествить его с процессами в мозге.

Идея Смарта заслуживает внимания. Что, собственно, он хочет сказать? Хочет ли он сказать, что ментальные состояния могут быть переопределены в терминах, которые не включают упоминания о чем-то нефизическом? Но что значит переопределены? Является ли данное им определение наличия ментального образа исчерпывающим? Если да и если у нас есть основания интерпретировать его формулу как описание процессов в мозге, то ментальных образов как особых данностей вообще нет. В одном месте статьи Смарт так и говорит: «В известном смысле таких вещей, как остаточные образы или чувственные данные, просто не существует, хотя и существует такая вещь, как опыт наличия образа, и этот опыт может быть косвенно охарактеризован на языке материальных объектов, но не на феноменологическом языке, поскольку его не существует»⁷³. Но вопрос в том, каков статус этого на первый взгляд сугубо элиминативистского утверждения Смарта? Присмотревшись к другим положениям его статьи, мы увидим, что это не более чем гипотеза.

В самом деле, если бы Смарт считал доказанным то, что ощущения, образы и другие подобные состояния исчерпывающе дефинируются на «языке материальных объектов», то он не мог бы рассматривать эпифеноменализм и теорию тождества как альтернативные метафизические допущения. Поэтому смысл его положений в другом: (1) есть некие данности, которые мы называем образами и т. п., (2) их наличие может быть выражено на физикалистском языке. Если оба пункта верны, то проще предположить, что физические события и эти ментальные данности — не разные вещи, а одна и та же вещь, хотя подобный тезис остается спекулятивной гипотезой.

Следующий шаг в развитии теории тождества был сделан Дэвидом Льюисом и Дэвидом Армстронгом. Они подхватили смартовскую идею «нейтрального» переопределения ментальных состояний и сформулировали новый аргумент в пользу теории тождества, мотивируя это тем, что одной бритвы Ок-кама, к которой прибегал Сمارт, было явно недостаточно. Аргумент состоял в следующем. Прежде всего трудно спорить с тем, что мозг опосредует человеческое поведение, выступая неким каузальным медиумом стимулов и реакций. Далее, можно попробовать показать, что все ментальные состояния можно интерпретировать в том же каузальном ключе: либо как причины поведения, как это имеет место в случае желаний, либо как результаты внешнего воздействия, как в случае восприятий или образов. И если это так (хотя так ли это — большой вопрос, и очевидности здесь нет, хотя и Льюис, и Армстронг поначалу считали, что это именно так и что тезис о каузальной действительности ментальных состояний может быть представлен как концептуальная истина, что неверно, так как, к примеру, для дефиниции желаний достаточно, чтобы они мыслились в качестве коррелятов — но не обязательно причин — поведения), то процессы в мозге и ментальные состояния играют одну и ту же роль, и вполне логично предположить, что они делают это именно в силу их онтологического тождества⁷⁴.

Каузальная теория ментальных состояний была выдвинута Льюисом в статье «Аргумент в пользу теории тождества» (1966)⁷⁵ и подробнейшим образом развита Армстронгом в книге «Материалистическая теория сознания» (1968). Исходя из сказанного выше, можно было бы думать, что эта каузальная теория является всего лишь дополнением концепций Плейса и Смарта. Однако ситуация более запутанна. Дело в том, что, в отличие от Смарта, Армстронг, похоже, не исключал, что ментальные состояния могут быть исчерпывающе определены в терминах их каузальной роли. И это означает, что его можно было бы причислить к элиминативистской традиции.

В самом деле, позиция Армстронга сводится к тому, что человеческий мозг реализует определенные каузальные схемы. Эти схемы можно было бы реализовать и на другом носителе. Поэтому они могут рассматриваться как нечто относительно самостоятельное и лишь контингентно тождественное процессам в мозге. Подобные каузальные схемы и есть ментальные состояния: «Ментальные состояния — это состояния личности, дефинируемые исключительно в терминах каузальных отношений»⁷⁶. Никаких других характеристик у этих ментальных состояний нет. Они лишены особого субъективного содержания и *de facto* совпадают с процессами в мозге.

Недостатки этой теории очевидны. Проводя концептуальный анализ и доказывая, что каждое ментальное состояние может быть истолковано в терминах его каузальной роли⁷⁷, Армстронг работает с этими состояниями как с интроспективными данностями. В самом деле, чтобы определять каузальную роль чего-то, он должен был принимать, что оно существует. Но затем, установив отношение этого нечего к стимулам и поведению, он неожиданно «нейтрализует» его, растворяя это нечего в каузальной роли, — и лишь позже возвращается к проблеме ее исполнителя, объявляя, что это процессы в мозге.

Понятно, что либо это просто паралогизм, либо Армстронг все же не элиминирует интроспективные данные, а неявно отождествляет их с процессами в мозге — именно исходя из того, что они играют одинаковые каузальные роли. В последнем случае его теорию действительно можно рассматривать как развитие идей Плейса и Смарта — тем более что, как мы видели, элиминативистские начала присутствовали и у них. Однако в любом случае изменения, внесенные Армстронгом, не сводятся к новым акцентам. Если Плейс, Фейгл и Сمارт говорили о тождестве процессов в мозге и таких особых состояний, как ощущения, образы и т. п., толкуя все другие ментальные состояния в райловском смысле, т. е. как поведенческие формы или диспозиции к поведению, то Армстронг утверждает, что подобное тождество распространяется на все ментальное, истолкованное в качестве суммы внутренних состояний, а именно «центральных состояний» нервной системы, т. е. мозга, играющих каузальные роли (при этом райловские диспозиции, которые казались чем-то бессубстратным и висящим в воздухе, обретают материальную основу)⁷⁸.

Акцентирование каузальной роли ментальных состояний сближает Армстронга с функционализмом — концепцией американского производства, которая допускает возможность интерпретации ментальных состояний в терминах осуществляемых ими функций, переводящих стимулы в поведенческие реакции⁷⁹. Не стоит, правда, забывать, что функционализм тяготеет к компьютерной метафоре сознания, чуждой Армстронгу⁸⁰. Это можно объяснить разным контекстом: создатель функционализма Хилари Патнэм учился у Куайна, виртуоза математической логики, тогда как учителя Армстронга в Оксфорде были далеки от подобных технических вещей.

Многие, кстати, воспринимали Патнэма как оппонента теории тождества. В 60-е годы он любил рассуждать о возможной множественной реализации ментальных состояний. Речь идет о ситуациях, когда, к примеру, боль и т. п. могут испытывать существа с другой биохимией, чем у человека⁸¹. Этот

довод подрывал тезис, что ментальные состояния тождественны определенным типам процессов в мозге. Но хотя он мог бы быть опасным для Плейса или Смарта, он не действует против Армстронга и Льюиса⁸², отводивших ментальным состояниям роль, сходную с той, которую приписывал им Патнэм.

Так или иначе, но рассуждения Патнэма ускорили трансформацию теории тождества. Одни ее сторонники отказались от тезиса о "типовом тождестве", т. е. о том, что ментальные состояния тождественны вполне определенным физическим процессам, допуская, что в каждом индивидуальном случае обнаруживается особое тождество. Другие перешли в лагерь функционализма, подчеркивая, что ментальные состояния совпадают скорее с функциональными, чем с собственно физическими характеристиками материальных систем.

Происходили и другие изменения. Американский философ Майкл Маккинси — один из противников теории тождества, атакующий ее с помощью семантических аргументов, — считает, что они сводились к тому, что позднейшие теоретики тождества, такие, как Джон Перри, автор «Знания, возможности и сознания» (2001)⁸³, в отличие от своих предшественников, признавали реальность субъективных ментальных состояний, которые, как уже отмечалось, часто именуются «квалиа», и тем не менее защищали тезис об их тождестве с процессами в мозге (Маккинси называет этот вариант теории тождества «квалиа-физикализмом»). Мы, однако, видели, что в этом вопросе Перри существенно не отступает от позиции ранних теоретиков тождества: он лишь более решительно избавляется от элиминативистских колебаний. Главные изменения, на мой взгляд, были в другом — в попытке разработать концептуальный аппарат, позволяющий объяснить уже не тождество, а различия, все же имеющиеся между ментальными состояниями и тождественными — согласно предположению теоретиков тождества — им физическими процессами.

И самый эффективный способ такого объяснения — допущение двойного когнитивного доступа к одним и тем же событиям в мозге. Это именно то, что несколькими страницами выше было названо эпистемологическим подходом к трактовке различения сознания и мозга. В наши дни выходит немало работ, авторы которых защищают именно такой взгляд, — одной из самых заметных в этом ряду стала, пожалуй, книга М. Вилманса «Понимание сознания» (2000)⁸⁴. Учитывая, что именно такой подход мог бы спасти от краха систему Сёрла, можно предположить, что он еще пополнит этот список.

Я, впрочем, очень сомневаюсь в продуктивности всех этих прошлых и будущих усилий. Дело в том, что теория тождеств-

ва, породившая эпистемологический подход к проблеме сознание—тело — шаткая конструкция. Причем, как мне представляется, для ее разрушения не требуется столь утонченной аргументации, к которой прибегали, в частности, Крипке, Чалмерс и Маккинси⁸⁵ и которую я не буду рассматривать⁸⁶ именно потому, что считаю ее избыточной.

Простой довод, сразу подрывающий теорию тождества, состоит в том, что ее главный тезис — если не интерпретировать его в элиминативистском смысле, когда он очевидно ложен, — попросту бессмыслен: это и есть та карусель иллюзий, на которой помещалась вся теория тождества⁸⁷. И бессмыслен этот тезис потому, что его нельзя верифицировать⁸⁸. В самом деле, как верифицировать утверждение, что некое ментальное состояние — скажем, чувство, возникающее при просмотре фильма *Lost in Translation*, — тождественно нейронному процессу в мозге? Инстанции, которые надо отождествить, кажутся совершенно различными. Теоретики тождества, разумеется, признают эту кажимость. Более того, мы уже знаем, что они говорили, будто тождество подобных феноменов имеет случайный, контингентный характер. Т. е. они допускали мыслимость, к примеру, такой ситуации, в которой данное чувство существовало бы вне связи с нейронными процессами. Но все это, по их мнению, не мешает тождеству⁸⁹. Утренняя звезда тоже случайно тождественна Вечерней звезде, и, мысля первую, мы необязательно мыслим и вторую. Проблема, однако, в том, что аналогия с Венерой не проходит.

Утверждение о тождестве Утренней и Вечерней звезды может быть легко верифицировано — достаточно проследить за движением этого небесного тела. Точно так же легко верифицировать, скажем, излюбленный тезис Плейса, что молния есть движение электрических зарядов, или положение, что на двух мониторах, к примеру, мы видим одну и ту же вещь. Замечу, что подтвердить последнее легче всего, если мы заметим корреляцию между этими изображениями, когда, к примеру, происходящее с вещью на одном мониторе происходит с ней и на другом. Таким образом, от корреляции мы можем заключать к тождеству — но только тогда, когда мы можем представить или показать, как скоррелированные явления возникают от одной-единственной вещи. В нашем случае мы могли бы проследить, откуда берутся данные для изображения, где находится видеокамера и т. п. Но никакими усилиями нельзя представить себе верификацию тезиса о тождестве качественных состояний и нейронных процессов. И это означает, что данный тезис просто не имеет смысла — или ложен.

Интересно, что могли бы ответить на это возражение сторонники теории тождества? Я поинтересовался на этот счет

у главного теоретика всего этого направления, Д. Армстронга, допуская, что он все же не является чистым элиминати-вистом⁹⁰. Его ответ был неожиданным, но логичным: «Должен признаться, я, по сути, никогда не рассматриваю вопрос о верификации! Думаю, мы уже ушли от принципа верификации»⁹¹. В самом деле, принято считать, что этот «венский» принцип оказался недостаточным, что Поппер, не взятый в Венский кружок и движимый благородным гневом, в свое время подорвал его⁹², а позднейшие попытки Айера и других авторов вернуть ему авторитет провалились благодаря И. Берлину и А. Черчу⁹³.

Но если непредвзято взглянуть на этот принцип и его критику, то мы увидим, что во многих случаях она не очень опасна. К примеру, говорили, что нельзя верифицировать общие утверждения, а это подрывает науку, которой должен был служить этот принцип. Очевидно, однако, что, хотя они не могут быть окончательно верифицированы, они верифицируемы в принципе. Скажем, верификация утверждения «всякое событие имеет причину» подразумевает, что какое бы событие мы ни наблюдали, нам удастся обнаружить у него устойчивые корреляты. Конечно, такая трактовка предполагает, что мы исходим из «слабого» понимания принципа верификации, согласно которому верифицируемые утверждения таковы, что из них — и их теоретического и эмпирического контекста — можно дедуцировать проверяемые наблюдения следствия, которые не дедуцируемы из одного контекста (согласно «сильной» трактовке верифицируемо то, что, наоборот, дедуцируется из набора эмпирических констатации)⁹⁴. Между тем «слабая» верификация, как утверждается, слаба настолько, что она вообще не позволяет различить бессмысленные и осмысленные утверждения⁹⁵. Но, хотя доказательство этого тезиса, приводимое, в частности, С. Сомсом⁹⁶, выглядит убедительно, оно имеет силу лишь для абстрактных положений и не затрагивает утверждения, констатирующие те или иные факты⁹⁷. А интересующий нас тезис о тождестве квалиа и процессов в мозге именно таков.

Совершенно неубедительным представляется и известное утверждение, что сам принцип верификации неверифицируем. Ведь этот принцип верифицируется так, как верифицируется любое аналитическое положение, тавтологическое в своей логической основе: он есть аналитическое следствие понятия истинности как соответствия мысли и предмета. Если мысль претендует на соответствие с предметом, то этот предмет должен быть дан. Если предмет не может быть дан — если мысль не верифицируема, — то ее утверждение не может быть истинным, а значит, оно с необходимостью ложно.

В любом случае мало кто считает, что принцип верификации вообще должен быть исключен из набора философских методов. Во многих случаях он прекрасно работает. И понять, что он работает, можно, задав вопрос: что имеется в виду под таким-то утверждением? Задав этот вопрос в нашем случае, мы увидим, что на него попросту нельзя ответить. В самом деле, что имеется в виду под тезисом о тождестве ментального состояния с каким-то процессом в мозге? Что это одна вещь? Но что значит одна, когда мы воспринимаем два феномена? Может быть, что это два аспекта одной вещи? Но тогда эта вещь — если «вещь» здесь не просто вербальная конструкция — должна мыслиться как что-то отличное и от ментального, и от физического, но ничего такого представить мы не в состоянии. Или, может быть, здесь просто имеется в виду тесная связь одного и другого, их корреляция? Но теоретики тождества отличают корреляцию, даже «совершенную», от отождествления.

Вот тут-то, возможно, и разгадка этого исторического казуса. Вспомним, что создатель теории тождества, У. Плейс, поначалу не различал эти вещи в интересующей нас области⁹⁸. Т. е. он считал, что отыскание нейронных коррелятов (говоря современным языком) ментальных состояний и является тем, что верифицирует их тождество. Смарт осознал, что верификации здесь не происходит, но вместо того, чтобы отбросить гипотезу Плейса как бессмысленную и ложную, он представил ее как метафизическую альтернативу эпифеноменализму⁹⁹, хотя эпифеноменализм как раз относится к числу осмысленных и верифицируемых допущений: условием его верификации является отыскание нейронных коррелятов ментальных состояний, от которых зависят последние и которые достаточны для объяснения поведения. Таким образом, на раннем этапе формирования теории тождества в результате, по сути, недоразумения (допущенного Фейглом и Смартом и проигнорированного Армстронгом, Льюисом и др.) произошло то, что философы стали выискивать аргументы в пользу не верифицируемой гипотезы, призванные перевесить верифицируемую гипотезу — заведомо провальное предприятие, ибо никакие аргументы не могут превратить ложный тезис в истинный.

Верифицируемая гипотеза (верифицируемая, конечно, еще не значит истинная, а лишь такая, которая могла бы в принципе оказаться истинной), с которой, как мы теперь видим, безуспешно сражались теоретики тождества, — это, еще раз напомним, эпифеноменализм, концепция, согласно которой ментальные состояния — как нечто реально отличное от нейронных процессов — порождаются ими, но не име-

ют никакого обратного влияния на эти процессы. И сейчас самое время вспомнить о том, что именно эпифеноменализм является самой реальной перспективой для Джона Сёрла, по которому читатели могли уже соскучиться и к анализу взглядов которого я возвращаюсь. Если не считать Сёрла некогерентным мыслителем, лишь но словах решающим все проблемы (ведь это была бы ирония судьбы: человек, призывающий снять с философских проблем словесную скорлупу, предлагал бы нам именно ее), придется признать его скрытым эпифено-меналистом. Ведь онтологическая нередуцируемость сознания - более важная вещь для него, чем утверждение каузальной действенности ментального: сам он говорит, что каузальная действенность ментальных состояний не выходит за пределы каузальной действенности нейронных процессов, лежащих в их основе.

И здесь, кстати говоря, удобный повод еще раз коснуться сёрловского «факта» порождения ментального нейронными процессами. Мы видели, что о фактичности и каузальности этого порождения можно говорить лишь в том смысле, что наличия нормально функционирующего человеческого мозга, по-видимому, достаточно для существования в этом мозге сознания. Напомним, что для решения проблемы сознание -тело, по Сёрлу, надо дополнить эту констатацию тезисом, что сознательные состояния, будучи онтологически нередуцируемы к нейронным процессам, тем не менее являются высокоуровневыми характеристиками этих процессов, т. е. тоже могут быть названы процессами в мозге и избавлены от дуалистических коннотаций.

Я уже отмечал проблематичность этого тезиса - и не будем к нему возвращаться. Вместо этого попробуем понять, избавлен ли от неясностей предшествующий ему каузальный постулат. Сразу надо подчеркнуть, что Сёрл проводит тут важное различие¹⁰⁰. В книгах конца 90-х «Тайна сознания» и «Сознание, язык и общество»¹⁰¹ он обозначает три вопроса: порождает ли мозг сознание? как мозг порождает сознание? и как может мозг порождать сознание?

Первый из этих вопросов допускает простой положительный ответ и составляет часть решения философской проблемы сознания — именно в силу простоты этого вопроса Сёрл считает элементарной и саму проблему сознание—тело, по крайней мере в этом ее аспекте. Второй вопрос, напротив, с его точки зрения, не имеет отношения к философии. Ответ на него должна давать нейронаука. Путем эмпирических разысканий ученые могут обнаружить нейронные корреляты и причины конкретных сознательных состояний. Сёрл одобрительно отзывается об исследованиях Ф. Крика и особенно

Дж. Эдельмана (который больше импонирует ему, так как трактует сознание как единую, холистскую среду, в отличие от «атомарного» понимания Крика), указывая, что они находятся на верном пути, хотя и в самом его начале. Наконец, третий вопрос — как мозг может порождать сознание? — объявляется Сёрлом все еще закрытым для решения: «Причина, по которой сознание кажется "тайной", в том, что у нас нет ясной идеи относительно того, как что-либо в мозге могло бы каузально породить сознательные состояния»¹⁰². Очевидно, что вопрос «Как мозг может порождать сознание?» может возникнуть лишь в ситуации, когда нам известно, что мозг порождает сознание, но неясно, почему это происходит. Но ведь это и есть наша «трудная проблема»!¹⁰³ Еще в начале книги я отмечал, что большинство современных авторов не сомневаются, что мозг порождает сознание, но, в отличие от своих предшественников, пытаются уточнить, при каких условиях и почему мозг порождает его. Сёрл же, как мы только что убедились, оставляет этот вопрос без ответа. Причем он не только не говорит, что мы пока не знаем ответа, но не утверждает и того, что его можно будет узнать. Ссылаясь на Макгинна, он замечает, что «могло бы оказаться так, что мы были бы просто не в состоянии объяснить это, что каузальное отношение мозга и сознания противится теоретическому объяснению, что проблема объяснения отношения сознания к мозгу выходит за пределы наших биологически ограниченных когнитивных способностей»¹⁰⁴. Однако позиция Макгинна для него тоже не более чем одна из опций. Возможно, говорит нам Сёрл, Макгинн прав и ответ на вопрос, как мозг может порождать сознание, в принципе не может быть получен, но не менее вероятен, считает он, тот вариант, что прогресс нейрона-уки и отыскание нейронных основ сознания все же позволит раскрыть соответствующие каузальные механизмы до такой степени, чтобы уяснить, почему над этими основами возвышаются субъективные состояния¹⁰⁵.

Итак, самый большой сюрприз ожидал нас, как выясняется, уже под конец нашего экскурса в философию Сёрла: он вообще отказывается отвечать на главный вопрос, решать «трудную проблему сознания». И теперь мы можем понять истоки неясностей, обнаруживаемых даже в тех частях его метафизики сознания, где он предлагает какие-то ответы. Не решив «трудную проблему сознания», нельзя надеяться и на четкое понимание его онтологического статуса.

Впрочем, несмотря на декларативный отказ Сёрла от рассмотрения «трудной проблемы», в его текстах все же можно найти намеки на то, в каком направлении он пытается нащупать ее решение. Чтобы разглядеть их, задумаемся, на что на-

целивает нас «трудный» вопрос о том, почему мозг порождает сознание. Мы уверены, что мозг достаточен для сознания, но не знаем, что именно делает его достаточным, не знаем, какое именно свойство мозга ответственно за это порождение. Решение «трудной проблемы сознания» позволило бы выявить это свойство и объяснить его действенность в качестве генератора сознания. Но как найти подходы к его выявлению? В XX веке многим казалось, что ответ на этот вопрос надо искать в сфере разработок по теме искусственного интеллекта¹⁰⁶. Успешное программное моделирование ряда когнитивных операций наводило на мысль, что мозг тоже может быть рассмотрен как компьютер. И если эта аналогия правомерна, то, возможно, сознание может быть истолковано как некая совокупность программ, инсталлированных в мозге, или как результат «программного» наполнения нейронных сетей.

Так и возникла гипотеза, что особенностью мозга, генерирующей сознание, является именно его вычислительная природа, способность десятков миллиардов его нейронов реализовать сложнейшие биологические и культурные алгоритмы, совокупность которых и является тем, что мы называем сознанием, Сёрл, однако, решительно выступил против программной метафоры сознания. В теперь уже почти легендарной статье 1980 г. «Сознания, мозги и программы» он предложил мысленный эксперимент, так называемый «Аргумент китайской комнаты», который был призван разрушить всякую аналогию между сознанием и программами.

Аргумент китайской комнаты — повествующий о человеке-компьютере, скрытом в помещении, обстановка которого позволяет создать впечатление его владения китайским языком для сторонних наблюдателей, хотя он не понимает по-китайски, — породил сотни откликов. Обсуждение не прекращается и поныне. Злые языки¹⁰⁷ говорят, что Сёрл позаимствовал его идею из одной статьи Нэда Блока, где тот рассуждал о моделировании работы мозга китайской нацией с радиопередатчиками¹⁰⁸. Сам Сёрл предлагает иную версию: «В первые дни когнитивной науки существовал некий фонд под названием "Слоан Фаундейшн", который давал людям деньги, чтобы они ездили по стране с лекциями. Нас в шутку величали Слоанскими Рейнджерами. Меня позвали в Йель прочитать лекцию в Лаборатории искусственного интеллекта. Я понятия не имел об искусственном интеллекте. Я купил книгу Роберта Шэнка и Роберта Эбелсона¹⁰⁹, из этой лаборатории, чтобы почитать в самолете. Я читал ее в самолете, и там шла речь о программах, понимающих рассказы. И я подумал: да это же смешно! Есть простое опровержение. Пусть это будет китайский рассказ и пусть я буду компьютером. Я все равно

не буду понимать рассказ. Так появился Аргумент китайской комнаты»¹¹⁰.

Отчет Сёрла весьма содержателен: он позволяет уяснить, почему его статья 1980 г. имела такой большой резонанс¹¹¹. В этой статье поначалу действительно идет речь лишь о частном случае программной эмуляции когнитивных способностей — и Сёрл без труда показывает, что при этом сохраняются громадные различия искусственного и человеческого интеллекта. Но там же Сёрл дает понять, что считает возможным обобщить свой анализ на все случаи подобной эмуляции¹¹². Таким образом ему удалось добиться мощного эффекта переноса очевидностей частного случая на его глобальный вывод о том, что естественный интеллект не работает как программа.

Чтобы разобраться, имел ли он право на такие обобщения, ненадолго углубимся в детали. Как мы только что слышали, Сёрл ознакомился с программами, обеспечивающими видимость понимания компьютером рассказов. Человеческое понимание всегда контекстуально, и если мы слышим рассказ о том, что человек зашел в ресторан и через час покинул его, оставив чаевые, то мы понимаем больше того, что буквально сказано. К примеру, если спросить нас, съел ли этот человек заказанную им сусину с икрой морского ежа (сам Сёрл скромно пишет о гамбургере, но я хотел бы приспособить его иллюстрации к российской специфике), то мы без колебаний ответим «да». Программа, эмулирующая понимание, должна отвечать на подобные вопросы. И в принципе такую программу можно написать, задавая контекст, на который будут накладываться вводимые данные. Но можно ли сказать, что компьютер, снабженный этой программой, действительно будет понимать рассказ? Сёрл убедительно показывает, что это не так. Но прежде он оговаривается, что подразумевает под программой алгоритм формальной обработки символов, без учета их возможной интерпретации — с такими синтаксическими программами и работают классические компьютеры.

Так вот, программу, понимающую некий рассказ, может «исполнить» и человек. При этом совершенно неважно, на каком языке будет этот рассказ, хоть на китайском, как и допускает Сёрл. Лишь бы этот человек знал — и это знание дает имеющаяся у него или заученная им программа, — какие именно символы выдавать в ответ на поступающие вопросы, в которых его интересуют опять же только символы, т. е. формальная сторона. Заучивший такую программу человек будет полным аналогом компьютера, эмулирующего понимание рассказа. Со стороны — наблюдателю, который находится вне Китайской комнаты, где сидит наш испытуемый, — будет казаться, что тот понимает рассказ. Но в действитель-

ности он не будет иметь никакого представления, о чем же идет речь в этом самом рассказе.

Отсюда следует, говорит Сёрл, что компьютер, выполняющий программу по эмуляции человеческого интеллекта, не может рассматриваться как аналог этого интеллекта, так как эта эмуляция происходит без какого бы то ни было реального понимания, характерного для нашего интеллекта. А это, в свою очередь, означает, что человеческий интеллект не есть программа, выполняемая на нейронном компьютере мозга. Тем самым, по мнению Сёрла, он опровергает концепцию «сильного искусственного интеллекта» (ИИ), согласно которой наличие сознания эквивалентно выполнению программы или необходимо связано с ним. Но все, на мой взгляд, могло бы повернуться иначе, если бы он не сделал одну ремарку -вначале относительно конкретной программы понимания рассказов, а затем тоже обобщенную на все случаи, — а именно что, отвергая «сильный ИИ», он является сторонником так называемого «слабого ИИ», т. е. считает, что действия человеческого интеллекта могут успешно симулироваться на компьютере¹¹³. В самом деле, почему бы и нет, рассуждает Сёрл. На компьютере можно смоделировать все, в том числе и работу интеллекта. Но моделирование человеческого интеллекта так же мало может породить в компьютере понимание, как моделирование грозы может вызвать дождь в этом самом компьютере.

Допущение возможности «слабого ИИ» кажется безобидным, пока мы не осознаем, что Сёрл, выходит, готов допустить создание таких, к примеру, программ, которые будут полностью эмулировать человеком понимание хотя бы того же китайского языка. И в позднейших вариантах изложения своего аргумента Сёрл говорит уже именно о глобальных программах и о том, что их выполнение на компьютере не дает с необходимостью эффекта реального понимания. С одной стороны, это логично: ведь его интересует не частный, а общий тезис о несходстве нашего осознанного понимания и искусственного интеллекта. Но проблема в том, что при таком обобщении Аргумент китайской комнаты становится уязвим для критики.

В самом деле, еще раз представим человека, запертого в комнате, заваленной горами китайских иероглифов, и снабженного программой, которая позволяет, по предположению, полностью эмулировать его знание китайского языка. Что же это за программа? Какого рода программа может подсказывать ему выбор иероглифов, создающий полную иллюзию осмысленного речевого поведения, т. е. позволяющий ему проходить знаменитый тест Тьюринга? Программа — это ал-

горитм, ставящий в соответствие данные на входе данным на выходе. Но какой алгоритм может поставить в соответствие данным на входе, означаящим, к примеру — о чем не знает испытуемый — такой вопрос: «Сколько сейчас времени?», такие данные на выходе, которые соответствовали бы критериям разумного речевого поведения, т. е. не позволили бы мне сразу сказать (если допустить, что мне известно, что в комнате есть часы), что это ответ машины? Никакая программа не может заранее прописать, в какое время я задам этот вопрос — если это не программа, кодирующая божественное провидение (и даже она не выручила бы, так как получается, что человек, задающий подобный вопрос, не мог заранее ознакомиться с ней, что нелепо, но необходимо допустить, так как в противном случае ничто не мешало бы ему задать свой вопрос в другое время¹¹⁴). Или представим, что человек, задающий вопросы, заранее принес в ту комнату какой-то предмет — к примеру, банку кока-колы — и поставил его на стол (это не противоречит условиям мысленного эксперимента Сёрла). И теперь он спрашивает: «Что вы видите на столе?» Невозможно вообразить, как человек в Китайской комнате смог бы ответить на этот вопрос при заданных Сёрлом условиях. Таким образом, чтобы осмысленно ответить на подобные вопросы, наш герой должен получить инструкцию: такие-то иероглифы означают то-то и то-то (например, «взгляни на часы», «смотри на стол», «опиши то, что видишь»), т. е. получить семантическую информацию¹¹⁵.

Но, получая ее, он начинает понимать китайский язык! Можно, правда, сказать, что Сёрл предвидел это возражение и ответил на него уже в разбираемой нами статье 1980 г. Дело в том, что мы должны помнить, что программы, о которых идет речь, в принципе написаны для машин, а не для человека: он лишь, так сказать, случайный гость в Китайской комнате. Поэтому очевидно, что для ответа на вопрос, подобный заданному выше — о том, «что вы видите на столе» и т. п., - программа должна содержать набор команд для какого-то механизма, снабженного телекамерой, и, исполняя эти команды, такой механизм повернет камеру в нужное место. Так что человек, на время инсталлирующий в себя данную программу, должен просто отдавать команды такому механизму при появлении тех или иных иероглифов, не имея никакого представления ни о смысле этих закодированных команд, ни о смысле символов на выходе этого механизма, трактуемых извне как ответ на заданный вопрос.

Одним словом, Сёрл уверен, что превращение Китайской комнаты в Китайского робота ничего не меняет¹¹⁶: человек, исполняющий программу понимания китайского языка, все

равно не понимает по-китайски. Однако с этим нельзя согласиться. Вдумаемся еще раз в условия нового мысленного эксперимента. Они не совпадают с условиями Китайской комнаты. В первом случае предполагалось, что человек полностью знаком с программой понимания китайского языка. Во втором - он не знает одного существенного компонента программы, а именно он не знает, почему символы, уходящие в устройство с телекамерой (при этом неважно, знает ли он, что там именно телекамера, он просто отправляет куда-то символы), порождают какие-то другие символы на выходе, которые он, в свою очередь, передает в качестве ответа на вопрос о том, что он видит вокруг себя.

Еще раз подчеркну: в самой программе эта связь должна быть прописана, в самой программе должно быть прописано, что такие-то данные телекамеры надо кодировать такими-то символами и т. п. Иными словами, в самой программе должен быть квазисемантический блок. Поскольку он неизвестен человеку, исполняющему программу, он, естественно, не понимает задаваемые ему вопросы. Но можно предположить, что их понимает весь робот, одним из механизмов которого оказывается человек внутри его. Кстати, одно из самых популярных в профессиональной философской среде возражений на Аргумент китайской комнаты состоит именно в том, что, хотя человек в ней не понимает китайского, его может понимать вся комната¹¹⁷. Обычно оно вызывает смех в студенческой аудитории. И справедливо: если взять Китайскую комнату в ее, скажем так, изначальной обстановке, то это предположение просто нелепо. Изначальная программа китайской комнаты не имела квазисемантического блока. Человек в этой комнате знал всю программу - крайне ограниченную, напомним, по своим эмулятивным возможностям¹¹⁸, - и рассмотрение комнаты в целом ничего не меняет. Но в механизированной китайской комнате — в китайском роботе — ситуация изменилась, в его программе есть «семантическая» часть, и она неизвестна человеку. Если бы он имел к ней доступ, то, исполняя эту программу, он стал бы понимать китайский язык, что, собственно, и требовалось доказать¹¹⁹. Впрочем, это не означает, что сам робот реально понимает китайский. Пока это просто гипотеза, которая не может быть опровергнута средствами Китайской комнаты Сёрла.

Так или иначе, но, уяснив все эти обстоятельства, мы поймем, что Аргумент китайской комнаты показывает не совсем то или даже совсем не то, что думает Сёрл. Этот аргумент не разрушает «сильный ИИ», потому что такое разрушение подразумевает доказательство реальной возможности точных программных эмуляторов человеческого интеллекта, притом

таких, которые не порождают осмысленное понимание: если эмулирующая его система заведомо не может проходить тест Тьюринга, нет ничего удивительного в том, что мы отказываем ей в понимании! Сёрл, разумеется, прав, различая синтаксис и семантику и говоря, что наличие чисто синтаксических эмуляторов не означало бы, что реализующие их системы что-то реально понимают. Но проблема в том, что, как мы только что убедились, он не доказывает невозможности того, что определенного рода синтаксические эмуляции с необходимостью генерируют по крайней мере некое подобие семантических отношений. Соответственно, он не доказывает и того, что в случае существования таких программ инсталлирующие их материальные системы не обладали бы реальным пониманием. Что же касается чисто синтаксических эмуляторов, не порождающих понимание (это Сёрл доказал), то мы увидели, что они невозможны, если говорить не о частных интеллектуальных задачах, а об интеллекте в целом. И мысленный эксперимент с Китайской комнатой помогает уяснить, почему дело обстоит именно так.

Именно поэтому Аргумент китайской комнаты может использоваться для опровержения «слабого ИИ», т. е. возможности точной синтаксической эмуляции разумного поведения человека, в том числе его речевого поведения. Но еще раз подчеркну: Сёрл не обладает ресурсами для ответа на вопрос, можно ли эмулировать такое поведение программами с «семантическими» блоками, т. е. такими программами, которые приспособлены для взаимодействия с реальностью и отражения ее характеристик, и если да, то будут ли машины, реализующие эти программы, обладать сознанием и, в частности, осознанным пониманием. Так что пора переходить к другим авторам.

И все же еще несколько слов о Сёрле. Поскольку сам он убежден, будто доказал, что физическая реализация таких программ, которые позволяют точно воспроизводить человеческое поведение, может и не порождать сознания, то человеческий мозг, вынужден заключить он, генерирует сознание не потому, что в нем реализуются какие бы то ни было программы. Но тогда почему же? Если главное в порождающем сознание мозге не нейронные «алгоритмы» — функциональные связи нейронных популяций, отдельных нейронов или чего-либо еще, — то, очевидно, решающую роль играют биохимические свойства мозга — такой довольно смелый ответ предложил Сёрл в той же статье 1980 г., где презентуется его Китайская комната¹²⁰.

Этот ответ вызвал много критики, и справедливо. Как вообще это можно помыслить, что сознание возникает вслед-

ствии определенных биохимических свойств мозга? Есть какие-то вещества или соединения, присутствующие в мозге и несущие сознательный заряд? Более странное решение «трудной проблемы сознания» сложно вообразить. Впрочем, Сёрл быстро отказался от него. Он стал говорить, что в мозге есть некое каузальное ядро, ответственное за возникновение сознания, но входят ли в него биохимические компоненты, пока неизвестно¹²¹, и тем самым пришел к той позиции, для которой характерен отказ от попыток философскими методами решить «трудную проблему»: она перекладывается на плечи будущей нейронауки. Как мы знаем, эта позиция была заявлена Сёрлом в ряде работ 90-х гг. Однако в опубликованной в 2004 г. книге «Сознание» он уточнил¹²², что, даже если нейронаука даст каузальное объяснение сознания, это не будет означать, что она объяснит все, что можно было бы объяснить касательно отношения сознания и мозга. Но полные объяснения — редкость и в физике. И он дает понять, что философы, требующие такого объяснения, поднимают псевдопроблему, спрашивают о том, о чем спрашивать бессмысленно.

Вдумавшись в эти пояснения Сёрла, мы увидим, что его отношение к «трудной проблеме» в середине 2000-х гг. нельзя суммировать той формулой, которая была уместна для работ, опубликованных в конце 90-х, а именно: «Мы не знаем, решается ли она». Из более поздних текстов Сёрла следует, что если она ставится как проблема демонстрации концептуальной необходимости определенного соотношения мозга и сознания (философ мог бы удовлетвориться только таким объяснением), то она вообще не может быть решена. Очевидно, таким образом, усиление контраста взглядов Сёрла с его оптимистичными установками 80-х гг. Но в любом случае его отказ от «биохимического» решения «трудной проблемы» представляется чем-то естественным или даже неизбежным: тот вариант не проходил по концептуальным соображениям¹²³. В конце концов, все вещества состоят из одних и тех же элементарных частиц, и разница между ними должна сводиться к различным отношениям, различным каузальным взаимодействиям между этими частицами. Т. е. так или иначе сознание должно генерироваться какими-то функциональными или каузальными отношениями между чем-то и чем-то¹²⁴. Но, каузальные схемы могут задаваться программными алгоритмами. Вспомним теперь, что исторически Сёрл начал свой путь к тезису о неразрешимости «трудной проблемы», оттолкнувшись от представления о некорректности компьютерной метафоры сознания. Но сейчас у нас может возникнуть чувство — подкрепляемое выявленными дефектами Китайской комнаты, — что он поспешил расстаться с этой метафорой.

Так, собственно, и считает наш следующий герой, Д. Деннет, главный защитник «программной» концепции сознания. Актуальность рассмотрения его идей очевидна и потому, что, в отличие от Сёрла, который, как мы видели, в итоге оставил «трудные» вопросы о сознании без решения, Деннет обещает объяснить сознание. «Объясненное сознание» — так и называется его главный труд, который можно без колебаний называть одним из шедевров современной мысли. Начнем с выяснения предыстории этой работы Деннета.

Примечания

¹ См., напр.: Мареева Е.В. Дж. Сёрл: старое и новое в понятии сознания // Философия сознания: история и современность. С. 216-231.

² Faigenbaum G. Conversations with John Searle. LibrosEnRed 2003. P. 170.

³ Searle J, R, The Mystery of Consciousness. L., 1998 (1997). P. 112. Деннет ответил Сёрлу в том же духе — *ibid.* P. 115.

⁴ *Ibid.* P. 152-157, 162.

⁵ Разумеется, речь идет о философии.

⁶ Falgenbaum G. Conversations with John Searle. P. 19. Зато, судя по другому (видео) интервью, в то время он много читал Джойса, Пруста и Кафку.

⁷ Searle J. R. Mind: A Brief Introduction. Oxford, 2004. P. 1-2.

⁸ *Ibid.* P. 111.

⁹ *Ibid.* P. 114.

¹⁰ См.: Faigenbaum G. Conversations with John Searle. P. 28, 31.

¹¹ Об этой оригинальной методике вспоминает А. Стролл, присутствовавший на семинарах Остина в Беркли в 1958 г. — см.: Stroll A. Twentieth-Century Analytic Philosophy. N. Y., 2000. P. 163-164.

¹² PI 360. Другой пример этого метода в действии — критические рассуждения Витгенштейна по поводу декартовского *cogito ergo sum* во время одной дискуссии 1947 г.: «Если кто-то скажет мне, глядя на небо: "Думаю, что пойдет дождь, следовательно я существую", — я его не пойму» — Monk R. Ludwig Wittgenstein: The Duty of Genius. N. Y., 1991 (1990). P. 496.

¹³ Разумеется, не все технические термины такого рода автоматически окажутся философскими. Они будут таковыми лишь в том случае, если реальные различия, раскрываемые с их помощью, имеют непосредственное отношение к какой-либо из философских проблем.

¹⁴ Не могу не упомянуть в этой связи замечательную книгу — Эдмондс Д., Айдиноу Дж. Кочерга Витгенштейна: История десяти минутного спора между двумя великими философами. М., 2004.

¹⁵ Интересно, что, когда Остин делал доклад в Кембридже, Витгенштейн тоже схватил кочергу. Но ему он не угрожал. См.: Эдмондс Д., Айдиноу Дж. Кочерга Витгенштейна. С. 206-207.

¹⁶ Faigenbaum G. Conversations with John Searle. P. 109.

¹⁷ Остин Дж. Три способа пролить чернила; философские работы. СПб., 2006. С. 210-211.

¹⁸ Хотя бы потому, что в ней используется специальная терминология, не почерпнутая из обыденного языка, - см.: Stroll A. Twentieth-Century Analytic Philosophy. P. 172.

¹⁹ В частности, сам термин «речевые акты» впервые появляется в статье Остина «Как говорить», вышедшей в 1953 г., — см.: Stroll A. Twentieth-Century Analytic Philosophy. P. 173. В это время Сёрл уже учился в Оксфорде — он оказался там в 1952 г., а покинул Англию в 1959-м.

²⁰ Faigenbaum G. Conversations with John Searle. P. 181-182.

²¹ См.: Strawson P. The Bounds of Sense: An Essay on Kant's Critique of Pure Reason. L., 1966. Ср.: Strawson P. Individuals: An Essay in Descriptive Metaphysics. L., 1964 (1959). P. 9-11.

²² Канта интересовал прежде всего факт существования априорных синтетических суждений.

²³ Wolff Chr. Der verminderten Gedanken von Gott, der Welt und der Seele des Menschen, auch alien Dingen uberhaupt, Anderer Theil. Andre Auflage. Frankfurt a. M., 1727. S. 411.

²⁴ Разумеется, сказанное не означает, что Сёрл испытал какое-либо влияние Вольфа. Речь идет просто об определенном — и любопытном — совпадении исследовательских программ.

²⁵ Searle J. R. Mind, Language, and Society. P. 8. Ср.: Searle J. R. Freedom and Neurobiology. P. 30.

²⁶ См., напр.: Остин Дж. Три способа пролить чернила, С. 207.

²⁷ Это произошло, вероятно, в середине 40-х гг.

²⁸ См.: Остин Дж. Три способа пролить чернила, С. 264.

²⁹ См.: Остин, Дж. Избранное. М., 1999. С. 15-135.

³⁰ Searle J. R. Intentionality: An Essay in the Philosophy of Mind. Cambridge, 1983. P. VII.

³¹ Searle J. R. Mind, Language, and Society. P. 149.

³² Ibid. P. 150.

³³ В «Сознании» (2004) он причисляет их, вместе с «субъективностью», к сущностным и неразрывно связанным друг с другом аспектам сознания. См.: Searle J. R. Mind. P. 136-137.

³⁴ Если бы меня попросили указать лучшую философскую книгу последних десятилетий, я, наверное, назвал бы именно эту работу (с удивительно блеклым и незапоминающимся названием).

³⁵ Searle J. R. Mind, Language, and Society. P. 51.

³⁶ Ibid. P. 52.

³⁷ Default positions. Сёрл намеренно использует этот компьютерный термин — ibid. P. 9.

³⁸ Х. Дрейфус, коллега Сёрла по Беркли, считает, что тот взял понятие Фона у Хайдеггера.

³⁹ В отрефлексированном виде такие допущения предстают как что-то ментальное. Но в фоновом режиме, считает Сёрл, они функционируют в качестве сугубо нейронных состояний. Сёрл вообще

близок к декартовскому отрицанию существования ментального бессознательного.

⁴⁰ Searle J. R. *Mind, Language, and Society*. P. 10.

⁴¹ Ibid. P. 45-55.

⁴² Ibid. P. 52.

⁴³ На всякий случай замечу, что мое понимание онтологических установок сформировалось независимо от Сёрла и первая вариация на данную тему была опубликована мной еще в 1997 г.

⁴⁴ Searle J. R. *Mind*. P. 112.

⁴⁵ Ibid. P. 114.

⁴⁶ См.: Green J. B., Palmer S. L. (eds) *In Search of the Soul*. P. 75-100.

⁴⁷ Searle J. R. *Mind, Language, and Society*. P. 49.

⁴⁸ Видимо, сознательные состояния лучше всего определять именно через понятие приватности, т. е. недоступности внешнему наблюдению. В принципе Сёрл не возражает против этого.

⁴⁹ Как он дан с позиции третьего лица.

⁵⁰ См., напр.: Searle J. R. *Mind*. P. 279-299.

⁵¹ Chalmers D. J. *The Conscious Mind: In Search of a Fundamental Theory*. N.Y., 1997(1996). P. 375; Kim J. *Mental causation in Searle's «biological naturalism»* // *Philosophy and Phenomenological Research* 55: 1(1995). P. 192.

⁵² См.: Searle J. R. *Why I am not a property dualist* // *Journal of Consciousness Studies*, 9: 12 (2002) P. 57-64. Но если вчитаться в эту статью, то мы увидим, что отличия сёрловской позиции от утверждений дуализма свойств сводятся прежде всего к терминологическим вариациям. Сёрл утверждает, что дуалисты свойств считают сознание чем-то находящимся «над» мозгом потому, что они придерживаются старой терминологии, предполагающей противоположность ментальных и физических свойств. Сёрл отрицает эту терминологию — и в этом, по сути, все отличие.

⁵³ Ср. формулировку в «Свободе и нейробиологии» (2007): «Все наши ментальные состояния каузально порождаются нейробиологическими процессами в мозге, и реализуются в мозге как его высокоуровневые или системные свойства» — Searle J. R. *Freedom and Neurobiology*. P. 40

⁵⁴ Searle J. R. *Mind, Language, and Society*. P. 62; Searle J. R. *Biological naturalism* // *The Blackwell Companion to Consciousness*. P. 334.

⁵⁵ Сам он не часто характеризует себя подобным образом — но только потому, что он употребляет термин «здравый смысл» скорее в обыденном, чем в философском ключе, — Searle J. R. *Mind, Language, and Society*. P. 11-12. Впрочем, в статье «Биологический натурализм» он прямо заявляет, что его позиция является выражением здравого смысла в его научном понимании.

⁵⁶ Searle J.R. *The Mystery of Consciousness*. P. 197.

⁵⁷ Searle J. R. *Mind*. P. 132. Кстати, дуализм — единственная из установок по умолчанию, которая решительно отвергалась Сёрлом. Напомним, однако, что в 2000-е годы он потеплел по отношению к этой установке, придя к выводу, что рациональное свободное действие предполагает допущение особого деятельного агента. К моменту написания этих строк Сёрл так и не выработал окончательной по-

зиции по поводу согласования этой идеи с другими своими теориями. Но я, кстати, совершенно не уверен, что дуализм является одной из установок по умолчанию.

⁵⁸ Ср.: Strawson G. Realistic monism: why physicalism entails pan-psychism // Strawson G. et al. *Consciousness and Its Place in Nature: Does Physicalism Entail Panpsychism?* Charlottesville, 2006. P. 13.

⁵⁹ Как выражается Т. Нагель — см.: Nagel T. *Other Minds: Critical Essays*. 1969-1994. N. Y., 1995. P. 106.

⁶⁰ Еще один вариант — допущение того, что как ментальные, так и нейронные процессы в полной мере определяют поведение (сверхдетерминация), — не проходит для Сёрла хотя бы потому, что он считает, что при отсутствии нейронных процессов ментальные не могли бы существовать.

⁶¹ Подробнее об этом принципе речь пойдет в 4 и 5 гл.

⁶² См., напр.: Searle J. R. *Consciousness, the brain and the connection principle: A reply* // *Philosophy and Phenomenological Research* 55: 1 (1995). P. 219.

⁶³ Searle J. R. *Mind*. P. 113-114.

⁶⁴ Ср.: Corcoran K. J. *The trouble with Searle's biological naturalism* // *Erkenntnis* 55: 3 (2001). P. 307-324.

⁶⁵ *Ibid.* P. 128.

⁶⁶ Яркую картину старого Оксфорда, в котором главное место занимала Традиция, рисует А. Айер. См.: Ayer A. J. *Part of My Life*. Oxford, 1978 (1977). P. 75-81. Историю о вылазке туда Витгенштейна занимательно рассказывает Монк — см.: Monk It. *Ludwig-Wittgenstein*. P. 496-497.

⁶⁷ Плейс считал, что первая формулировка этой теории была предложена психологом Э. Борингом в 1933 г. — см.: Place U. T. *Identifying the Mind: Selected Papers of U.T. Place*. N. Y., 2004. P. 89. Любопытно, однако, что Г. Фейгл еще в 1934 г. утверждал, что Боринг был лишь одним из многих, кто отстаивал теорию тождества. См.: Feigl G. *Logical analysis of the psychophysical problem: a contribution of the new positivism* // *Philosophy of Science* 1: 4 (1934). P. 437. А Дж. Пратт, из статьи которого, вышедшей в 1936 г., также следует, что уже в середине 30-х гг. идея тождества ментального и физического была чем-то общеизвестным, вообще прямо не связывал эти идеи с именем Боринга (хоть и упоминал его в своей статье) — Pratt J. B. *The present status of the mind-body problem* // *The Philosophical Review* 45: 2 (1936). P. 144-166. О Г. Т. Фехнере и других более ранних предшественниках теории тождества подробно пишет М. Хайдельбергер — см.: Heidelberger M. *Wie das Leib-Seele Problem in den logischen Empirismus kam* // *E-Journal Philosophie der Psychologie* 2 (2005). <http://www.jp.philo.at/texte/HeidelbergerMl.pdf>. Так или иначе, но наличие многочисленных предшественников Плейса, Фейгла и других не означает, что они не могут рассматриваться как основоположники нового направления: это еще один пример, подтверждающий то, что история традиции не совпадает с историей идей — их не стоит смешивать.

⁶⁸ В этом тексте Фейгл выступил как сторонник теории тождества. В более ранний период он склонялся, скорее, к элиминативизму. Он полагал, что ментальные термины могут быть без остатка переведе-

ны на физикалистский язык. См.: Feigl G. Logical analysis of the psychophysical problem: a contribution of the new positivism // *Philosophy of Science* 1: 4 (1934). P. 420-445.

⁶⁹ Smart J. Sensations and brain processes // *The Nature of Mind*, ed. by D. Rosenthal. N.Y., 1991. P. 175.

⁷⁰ Этот термин заимствован Смарттом из упомянутого эссе Фейгла 1958 г., в котором, кстати, встречаются и другие, как оказалось, удачные термины — «квалиа», «нейронные корреляты» и т. п. Впрочем, только что упомянутые термины придумал не Фейгл. А об отыскании «нервных коррелятов» конкретных ментальных состояний как «насушной проблеме» говорили еще в начале XX в. См., напр.: Meyer M. The present status of the problem of the relation between mind and body // *The Journal of Philosophy, Psychology and Scientific Methods* 9: 14 (1912). P. 371.

⁷¹ У Фейгла в «Ментальном и физическом» эту линию обнаружить гораздо труднее.

⁷² Smart J. Sensations and brain processes. P. 172.

⁷³ Ibid. P. 173-174.

⁷⁴ См.: Armstrong D. A Materialist Theory of the Mind. Rev. ed. L., 1993. P. XIV. Впрочем, новизна этого аргумента относительна — нечто похожее мы встречаем уже у Лукреция. См.: Montero B. Varieties of causal closure // *Physicalism and Mental Causation: The Metaphysics of Mind and Action*, ed. by S. Walter and H.-D. Heckmann. Exeter, 2003. P. 181. И на Армстронге история этого аргумента не заканчивается. Вот, к примеру, одна из его новых версий: Humphrey N. How to solve the mind-body problem // *Journal of Consciousness Studies* 7 (2000). P. 5-20.

⁷⁵ См. об этой статье: Армстронг Д.М. Льюис и теория тождества // *Историко-философский альманах: Вып. 2. М., 2007. С. 227-231.*

⁷⁶ Armstrong D. A Materialist Theory of the Mind. P. 356; см. также p. 79.

⁷⁷ См.: Armstrong D. A Materialist Theory of the Mind. P. 129-351. Позже Армстронг усомнился, что в данном случае речь идет именно о концептуальном анализе, допуская возможность того, что его результаты могут быть проинтерпретированы как эмпирическая теория о работе сознания — см.: Armstrong D. *The Mind-Body Problem: An Opinionated Introduction*. Oxford, 1999. P. 88.

⁷⁸ Смарт согласился с Армстронгом, Плейс — нет. Он считал, что диспозиции зависят от состояний мозга, но не тождественны им — см.: Place U. T. *Identifying the Mind*. P. 97.

⁷⁹ Армстронг не возражает против именованной своей теории «каузальным функционализмом» — Armstrong D. *The Mind-Body Problem*?. 81.

⁸⁰ Он прибегает к ней лишь в исключительных случаях.

⁸¹ Putnam H. Psychological predicates // *Art, Mind, and Religion*, ed. by W. Capitan and D. Merrill. Pittsburgh, 1967, P. 37-48.

⁸² Ср.: Lewis D. Mad pain and Martian pain // *The Nature of Mind*. P. 229-234.

⁸³ Перри, кстати говоря, является соведущим еженедельного радишоу *Philosophy Talk*. Эта великолепная передача уже несколько лет выходит в Сан-Франциско и доступна в Интернете.

⁸⁴ Вилманс называет свою теорию «рефлексивным монизмом».

⁸⁵ Я всегда считал Майкла Маккинси, которого лично знаю с начала 90-х гг. (он читал тогда на нашем факультете в МГУ спецкурс по философии языка), образцовым аналитическим философом. Поэтому я обрадовался возможности еще раз пообщаться с ним в 2006 г. Просмотрев сейчас свои записи об этой беседе, я подумал, что они могут оказаться небезынтересными и создать представление о «кухне» аналитической философии. — 3 мая 2006 года, ТРАМ, встреча с Майклом Маккинси, в которой участвовали также Дмитрий Волков и жена М. Памела. Разговор начался в 12:00 на Пушкинской площади и продолжался до 15:30, последние 10 минут опять же на улице и площади. По пути в ТРАМ (к сожалению, этот замечательный и очень уютный ресторан с тех пор закрылся) рассуждали о Бродвейских постановках, которые видели недавно, а также о Деннете и Блэкмор. В самом ресторане после краткого политического вступления — Памела и М. сказали все, что они думают о Буше, и обсудили возможных кандидатов на пост президента {М. заметил, что если президентом вдруг станет брат Буша, то он эмигрирует в Канаду) — разговор перешел на философию. Я спросил, что М. думает о современном состоянии аналитической философии. Он сказал, что оно не очень радужно и связано это с тем, что философия языка, которая составляет, по его мнению, основу аналитической философии, переживает сложные времена. Куда-то в прошлое ушла точность мысли, на ее месте появилась вседозволенность. Впрочем, в последнее время видны некоторые признаки улучшения. Восстанавливается интерес в исследованиям языка. Назвал некоторые новые имена — в основном его студенты или коллеги. Я сказал, что Чалмерс тоже еще молод. Чалмерс вызывает у него уважение. Он поддерживает классическую аналитическую философию в эти «непростые дни» (я совершенно не согласен с такими оценками). М. переписывался с Чалмерсом. Отчасти солидаризировался, отчасти высказывал свои критические замечания, Чалмерс в ответ просто повторял свою позицию. Критикует мысленный эксперимент Чалмерса с зомби. Согласно концепции Чалмерса, получается, что зомби, ударивший свою ногу и прыгающий «от боли», не чувствует боль. Считает это абсурдом, хотя в ответ на то, что это просто следует из дефиниции зомби, ничего убедительного не возражает. Чалмерс спрашивал Маккинси, когда он наконец опубликует свою книгу. Над этой книгой М. работает уже более 20 лет, но все время изменяет свою позицию, и поэтому приходится откладывать публикацию. В отличие от Чалмерса Деннет не вызывает у него симпатии. Считает его «пиарщиком» (такой взгляд на Деннета весьма распространен в США в среде академических философов). Подмена тезисов в «Объясненном сознании» (ссылался на рецензию Н. Блока), упускает само сознание. Впрочем, с работами Деннета знаком не очень хорошо. Блэкмор вообще не знает. Отношение к Сёрлу чуть более теплое. Читал его «Новое открытие сознания», поддерживает Аргумент китайской комнаты и, соответственно, критикует функционализм за упущение квалиа. Но представление о решении Сёрлом проблемы сознание—тело самое общее. «А в чем его позиция, собственно, состоит?» Недоумение по поводу того, что Сёрл считает «трудную про-

блему» легкой. Конечно, ментальное порождается мозгом, но почему — вот ведь вопрос. Возможно, считает М., ответ на него будет получен нейронаукой — но заранее сказать нельзя. Симпатизирует позиции Макгинна (сообщил, что тот только что переехал во Флориду), но, с другой стороны, считает ее, возможно, слишком резкой. Что касается Фодора, то жалеет, что мало его читал. Восхищается Хомским — «легендарная личность», хотя политические взгляды последнего слишком левые даже для М. Относительно универсальной грамматики и врожденных структур полагает, что они существуют, но насколько они детализированы, сказать затрудняется. Считает интересной и модульную теорию сознания, но не решается заявлять о ней ничего конкретного. Но Линкера уважает. Уважает и Дж. Перри, спору с которым посвящена его лекция в Москве. Уважая его, изумляется, как он мог прийти к позиции квалиа-физикализма. О Рорти отзывается с неодобрением, удивляется, когда узнает, что Р. объявляет себя последователем Дэвидсона (считает Рорти постмодернистом, с интересом слушает о том, что тот дружил с Дер-рида). Дэвидсон, впрочем, тоже не вызывает у него пиетета, как и учитель Дэвидсона Куайн. Куайн был замечательным логиком, писал прекрасные, исключительно ясные работы (а вот учитель Ку-айна Карнап сильно ошибался в своих модальных рассуждениях, опираясь на ошибочную теорию смысла Фреге). Но потом, начиная с 50-х годов, Куайн стал откровенно «булшитить». Его статью о двух догмах буквально разгромили вскоре после выхода, в частности Стросон. Его критика различения аналитических и синтетических Предложений — полная ерунда. Не принимает ее. Впрочем, и оксфордская философия обыденного языка, и вообще оксфордский стиль философии тоже не привлекают его. Они производят полусырые продукты. Удивляется, когда узнает, что Остин говорил Сёрлу, что его тексты «полусырые». Замечает, что, в отличие от других оксфордцев, Остин, однако, сам не публиковал такие работы. Ему ближе Кембридж, ценит Рассела, Витгенштейна, особенно позднего (хотя и считает, что по сути тот ошибался) и, конечно же, Мура. Мур (как и Фреге, подлинный зачинатель аналитической философии — хотя ему не хватало философского применения его логических учений, отсюда значение Рассела, который, впрочем, был более поверхностным мыслителем) — главный авторитет. На замечание о том, что Мур не был сильным логиком, отвечает, что был — еще каким (ссылается на статью Мура о внутренних и внешних отношениях), хотя и не всегда это показывал. Опирался на обыденные интуиции, был точным и ясным. Маккинси предлагает делать то же самое. Философия — априорная дисциплина, базирующаяся на языковых интуициях и проясняющая их. Признает, что некоторые традиционные философские проблемы бессмысленны. К примеру, бессмысленна проблема смысла жизни. Но проблема сознание—тело не такова. Это реальная проблема, хотя решить ее можно на уровне семантики, решить в том смысле, что показать ошибочность редуктивного материализма и истинность нередуктивного материализма. Нередуктивный материализм понимает как учение, согласно которому ментальные события — это события в мозге, т. е. в каком-то смысле они совпадают с физическими процессами в нем, но при

этом они наделены особыми, нефизическими свойствами. Не слишком ясно. Проблему ментальной каузальности считает трудной, но довольно четко формулирует свою позицию: ментальные состояния позволяют объяснять поведение, но не являются его производящими причинами (как известно, подобное решение было, в частности, предложено Ф. Джексоном и П. Петитом). При этом отрицает, что такая позиция имеет эпифеноменалистский характер. Подробно излагает свои взгляды по принципиальным вопросам, от референции до трактовки ментальных свойств. Считает ключевой модальную логику. Отсылает к Муру, впервые в современной философии заговорившему на эту тему, замечает, что средневековые авторы были хорошо знакомы с модальной логикой, ссылаясь на некоторых медиевистов, которые проводили соответствующие изыскания (вообще признает пользу изучения истории философии, звучат имена Аристотеля, Канта, Юма, Декарта — себя не раз именует «картезианцем» — и Оккама). Говорит о своем опровержении зкетернлиз-ма Патнэма (которого обвиняет в грубых ошибках) и о своей полемике с Кринке (восхищается его дарованиями, рассказывает о неопубликованных текстах Кринке, которых очень много), о том, что Кринке в частной беседе признавал справедливость возражений М., но официально не отказался от своей каузальной теории референции. М. считает ее ложной, а свою альтернативную теорию — действительно более удачной альтернативой. Шутит, что считает себя лучшим философом языка, на втором месте — Сомс. Размышляет о полемике Сомса с Чалмерсом по поводу «двумерной семантики». Говорит, что занимает позицию между Сомсом и Чалмерсом. Указывает на ошибки чалмерсонской теории двух уровней значений, связанные с непониманием Чалмерсом некоторых положений доктрины возможных миров. Приводит интересные примеры. Всплывает, конечно, и "Утренняя с Вечерней звездой. Возвращаемся — уже на улице — к физикализму. Не согласен с тезисом о том, что квалиа-физикализм может быть найден уже у Армстронга. Теория Армстронга совпадает с концепцией Льюиса, это ошибочные функционалистские доктрины — частично функционализм может, впрочем, и работать... Через два дня, 5 мая, была лекция «Проблемы квалиа-физикализма» у нас в корпусе, в 1161. Перед лекцией рассказывал о Детройте, где живет. Небольшой даунтаун, с полупустыми небоскребами. Впрочем, в последнее время Детройт опять переживает подъем, хотя новые фирмы и не базируются в центре города. Кафедра философии в университете Уэйна, где он работает, небольшая, всего восемь человек, так как университетское начальство не очень серьезно относится к философии... Присутствовало человек тридцать, в том числе Н.С. Юлина, С.А., Диана и Саша, Н. Гарнце-ва, Косилова, Бугай, Мельников, Селиванов и многие другие. Читал М. по своим записям, которые были розданы слушателям. Отвечал на вопросы — всего их было пять, два из них задал Д. Волков — более живо. Главная идея лекции — квалиа-физикализм ложен, так как признание его истинным означало бы, что высказывания людей о своих ментальных состояниях должны логически имплицировать высказывания о состояниях мозга, о которых они, однако, могут вообще не иметь никакого представления. Основывал этот вывод на

различении пропозиционального и лингвистического значения. Пропозициональное определяется референтами, лингвистическое зависит от общеустановленных правил словоупотребления. Скажем, высказывание «я поливаю луг» всегда то же в лингвистическом но различно в пропозициональном смысле. Дальше он перекидывал мост к квалиа. Много также рассуждал в этой связи о понятии возможности и о разных ее видах. Рассуждения казались довольно туманными. После лекции я спросил, почему обсуждение вопроса о тождестве ментального и физического обычно получается таким трудным. Потому что это трудная тема, ответил М. Проблема квалиа очень сложна, так как непросто схватить их природу. Решать «трудную проблему» надо с двух сторон — нейронауки и семантики. Но семантический анализ крайне непросто. В будущем, возможно, эти линии будут все ближе сходить, и в конечном итоге это даст возможность все же разрешить проблему сознание—тело... После лекции встреча с М. в Патио на Л-ком с участием Памелы, Дианы, Саши и С. Маккинси и Памела были со всеми около двух часов. Обсуждали общие темы. М. рассказывал о своих песенных увлечениях и сам пел. Но не только. Пара философских тем была затронута. Еще по пути я высказал мнение о том, как можно более простым способом разгромить теорию тождества. Можно сделать это в стиле раннего (точнее, «среднего») Витгенштейна, показав, что тезис о тождестве ментального и физического просто лишен смысла, так как у нас нет никакой возможности верифицировать его. В ответ Маккинси сказал, что верифицируемость не имеет решающего значения и что в данном случае вышеупомянутый тезис может рассматриваться как «лучшее объяснение» корреляции ментального и физического (в этой связи речь зашла о знаменитой статье Смарта, которую Маккинси считает самым важным событием в истории аналитической философии XX века — я забыл спросить его, общался ли он со Смартом во время своей недавней поездки в Австралию). Я заметил, что, хотя как всеобщий принцип он и в самом деле не работает, в ряде случаев он может применяться. М. ответил, что тогда эти ограничительные условия должны быть сформулированы. В связи с рассуждениями о бессмысленности неverified предложений Маккинси усомнился в самой корректности понятия бессмысленности. Все так называемые бессмысленные предложения на деле попросту ложны. Например, фраза Хомского «зеленые идеи яростно спят» не бессмысленна, а ложна. Я ответил, что, на мой взгляд, не имеет большого значения, бессмысленна фраза или ложна. Пусть тезис о тождестве будет ложным — еще лучше. Ведь именно это и пытался показать Маккинси на своей лекции. В общем, в итоге вопрос остался открытым... Обсуждались и знакомства М. с разными влиятельными философами. Он недавно общался с Чалмерсом, когда-то контактировал с Сёрлом, а с Деннетом никогда не пересекался.

⁸⁶ Хотя, если кому-то покажется неубедительным мой довод, я отсылаю его к работам упомянутых авторов, где изложены семантические аргументы против теории тождества. См., напр.: McKinsey M. Refutation of qualia-physicalism // *Situating Semantics: Essays on the Philosophy of John Perry*, ed. by M. O'Rourke and Corey Washington. Cambridge MA, 2007. P. 469-498. Кроме того, в пя-

той главе будет предложен еще один независимый аргумент против теории тождества.

⁸⁷ Сёрл, кстати, обращает внимание на фундаментальную амбивалентность ранней теории тождества. См.: Searle J. R. *The Rediscovery of the Mind*. Cambridge MA, 1994(1992). P. 27-29.

⁸⁸ Ср.: Stroll A. *Twentieth-Century Analytic Philosophy*. P. 257-256. Правда, он менее категоричен. Еще более осторожно высказывался Дж. Грей, также рассуждавший на близкую тему — см.: Gray J. *Consciousness: Creeping Up on the Hard Problem*. N. Y., 2006 (2004). P. 111. В сходном ключе критиковал прототеорию тождества и Р. Карнап в «Логическом построении мира» (1928) — см.: Heldetberger M. *Wie das Leib-Seele Problem in den logischen Empirismus kam*. S. 19-20. Нельзя не напомнить в этой связи, что один из создателей теории тождества, Г. Фейгл, входил в Венский кружок, одним из лидеров которого был Карнап. Впрочем, по воспоминаниям самого Фейгла, в кружке его называли «метафизиком» (*ibid.* S. 21). Так что в его отступничестве нет ничего удивительного — он и в самом деле зачастую противостоял Карнапу. С другой стороны, в 30-е годы Фейгл во многом поддерживал его физикализм.

⁸⁹ После работ Крипке многие стали говорить, что всякое тождество необходимо, — и это добавило проблем теоретикам тождества. Но в данном случае это не имеет большого значения.

⁹⁰ В конце концов Армстронг полемизирует с элиминативистами. См.: Armstrong D. *The Mind-Body Problem*. P. 91-99.

⁹¹ Письмо от 6 января 2006 г.

⁹² Все началось с этой статьи 1933 года: Поппер К. Критерий эмпирического характера теоретических систем // Журнал «Erkenntnis» («Познание»). Избранное. М., 2006. С. 335-337.

⁹³ Показательно, что Айер признал свое поражение — см.: Ayer A. J. *The Central Questions of Philosophy*. N. Y., 1976 (1973). P. 27.

⁹⁴ См.: Soames S. *Philosophical Analysis in the Twentieth Century*. V. 1. Princeton, 2005 (2003). P. 277-291. Ср.: Ayer A. J. *Language, Truth and Logic*. 2 ed. N. Y., 1971(1946). P. 13-17.

⁹⁵ Подобным образом можно критиковать и попперовский принцип фальсификации.

⁹⁶ Soames S. *Philosophical Analysis in the Twentieth Century*. V. 1. P. 289-291.

⁹⁷ Иными словами, оно не затрагивает «прямо верифицируемые утверждения» («S прямо верифицируемо, если (a) S — утверждение наблюдения или если (b) S само по себе или с одним или множеством утверждений наблюдения P, Q, R,... логически влечет утверждение наблюдения, которое не следует из одних лишь P, Q, R,...») и показывает лишь неадекватность понятия «косвенно верифицируемого утверждения» («S косвенно верифицируемо, если (a) S само по себе или в сочетании с другими посылками P, Q, R, ... логически влечет прямо верифицируемое утверждение D, которое не следует из одних лишь P, Q, R,...; и (b) другие посылки P, Q, R, ... либо аналитичны, либо прямо верифицируемы, либо относительно них можно независимо показать, что они косвенно верифицируемы») — Soames S. *Philosophical Analysis in the Twentieth Century*. V. 1. P. 286). В самом деле, Сомс, следующий Берлину и Черчу, рассуждает так: до-

пустим, что *S* — это произвольное утверждение, а *R*, не-*R* и *O* — независимые утверждения наблюдения. Сочетание *S* с кондиционалом «если *S* или *R*, то *O*» влечет *O*. Этот кондиционал прямо верифицируем (см. определение), «так как его сочетание с *R* влечет утверждение наблюдения *O*, которое не выводимо из одного лишь *R*». Соответственно, если утверждение наблюдения *O* не следует из одного лишь прямо верифицируемого кондиционала «если *S* или *R*, то *O*», а следует из него в сочетании с *S*, то *S* должно быть признано косвенно верифицируемым и осмысленным (см. выше). Если же *O* следует из одного лишь кондиционала «если *S* или *R*, то *O*», то путем логических преобразований несложно убедиться, что оно следует из «Не-*S* & не-*R*», что, в свою очередь, означает, что не-*S* прямо верифицируемо, а это, как показывает Сомс ранее, значит, что его отрицание, а именно *S*, должно быть косвенно верифицируемо. Эти варианты исчерпывают все случаи. Таким образом, при данных критериях верифицируемости любое предложение оказывается верифицируемым (ibid. P. 290-291). На основании этого Сомс делает вывод о крахе принципа верификации. Однако он не следует из его доказательства. Еще раз подчеркну, что из него следует лишь тезис о неэффективности понятия косвенной верифицируемое™. Но это еще не повод полностью отбрасывать данный принцип и понятие прямой верифицируемости. Для того чтобы оправдать такой шаг, Соме должен был бы показать, что произвольно взятое утверждение можно истолковать в качестве прямо верифицируемого. Но неясно, как сделать это.

⁹⁸ По сути, он так и не изменил эту позицию. В одной из своих поздних статей он озаглавил ее заключительный параграф так: «Совершенная корреляция — это тождество» — см.: Place V. Token — versus type-identity physicalism // *Antropology and Philosophy* 3: 2 (1999). P. 21-31. Вместе с тем Плейс стал уточнять, что корреляция ментальных и физических состояний прямо указывает на их тождество только потому, что эпифеноменализм, который мог бы быть альтернативной интерпретацией данной корреляции, является некогерентной позицией: он не позволяет объяснять возможность знания о субъективных состояниях, существование которых он предполагает — см.: Place V. T. *Identifying the Mind*. P. 79-80. К этому вопросу мы еще вернемся в пятой главе этой книги. Пока хочу лишь заметить, что, соглашаясь с выводом Плейса о некогерентности эпи-феноменализма, я не могу принять его тезис, что в таком случае нам остается лишь сделать эмпирический выбор между интеракционизмом и теорией тождества (ibid. P. 80). Не говоря уже о том, что теория тождества едва ли осмысленна, аргументы, опровергающие эпифеноменализм, могут, как мы увидим, быть разрушительными и для теории тождества.

⁹⁹ Впоследствии Смарт изменил свое мнение и стал говорить, что это все-таки научная гипотеза, как и считал Плейс. Но это не значит, что их взгляды сблизились. Плейс полагал, что теория тождества верифицируется при отыскании нейронных коррелятов сознания, а Смарт утверждал, что она научна, так как выбор в ее пользу осуществляется на основе таких «научных принципов», как «бритва Оккама и соображения простоты» — см.: Smart J. J. C. *Postscript*

(1995) // *Philosophy of Mind: Contemporary Readings*, ed. by T. O'Connor, D. Robb. N. Y., 2003. P. 135.

¹⁰⁰ См.: Searle J. R. *The Mystery of Consciousness*. P. 191-193. Ср.: Searle J. R. *Mind, Language, and Society*. P. 51.

¹⁰¹ См. также: Searle J. R. *Consciousness and Language*. Cambridge, 2002. P. 10, 24.

¹⁰² Searle J. R. *The Mystery of Consciousness*. P. 193.

¹⁰³ Соответствие третьего вопроса Сёрла «трудной проблеме» Чалмерса не вызывает сомнений. В «Сознающем уме» (1996) Чалмерс формулирует ее так: «Как физическая система мол-сет породить сознательный опыт?» — Chalmers D. J. *The Conscious Mind*. P. 25. Сходство очевидно, но это не значит, что Сёрл использовал формулировку Чалмерса. Она есть уже в статье «Проблема сознания», впервые опубликованной в 1993 г. См.: Searle J. R. *Consciousness and Language*. P. 10.

¹⁰⁴ Searle J. R. *The Mystery of Consciousness*. P. 197-198.

¹⁰⁵ Ibid. P. 197-198; Searle J. R. *Mind*, P. 146. Ср.: Searle J. R. *Freedom and Neurobiology*. P. 32. В «Свободе и нейробиологии» (2007), как, к примеру, и в статье «Проблема сознания» (1993), Сёрл более оптимистичен по поводу перспектив научного решения проблемы сознания. Но в 2000-е годы он уже не считает, что ее научное решение закроет все философские вопросы.

¹⁰⁶ В Новое время то, что сейчас именуется «искусственным интеллектом» (термин «artificial intelligence» введен Джоном Маккар-ти в середине XX века), иногда называли «неодушевленным разумом» (inanimate reason). См.: Standage T. *Der Turke*. Berlin, 2005. S. 60.

¹⁰⁷ См., напр.: Pinker S. *How the Mind Works*. P. 93.

¹⁰⁸ Block N. *Troubles with functionalism* // *Minnesota Studies in the Philosophy of Science*, 1978, 9. P. 261-325. В действительности мысленный эксперимент Блока имеет другую структуру, чем аргумент Сёрла. Блок просто обращается к интуициям здравого смысла: китайская нация не может быть наделена каким-то отдельным от индивидуальных сознаний сознанием. Сёрл же предлагает человеку реально встать на место компьютера — и только после этого делать выводы.

¹⁰⁹ Вот книга, которую купил Сёрл: Schank R., Abelson R. *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structure*. Hillsdale, 1977.

¹¹⁰ Faigenbaum G. *Conversations with John Searle*, P. 54.

¹¹¹ Некоторые — в шутку, а может, и не совсем в шутку — даже предлагали переопределить когнитивную науку как «исследовательскую программу, направленную на доказательство ложности сёрловского Аргумента китайской комнаты» - см.: Hamad S. *Minds, machines, and Searle 2: What's right and wrong about the Chinese Room Argument* // *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence*, ed. by J. Preston and M. Bishop, N. Y., 2002. P. 295.

¹¹² Я использую термин «эмуляция» как синоним термина «симуляция», применяемого Сёрлом.

¹¹³ Searle J.R. *The Mystery of Consciousness*. P. 17.

¹¹⁴ Кстати, подобные программы неявно исключаются Тьюрингом, по образцу знаменитого теста которого построена Китайская комната. См.: Turing A, Computing machinery and intelligence. *Mind* 59 (1950). P. 433-460. Это очевидно из финала 6 параграфа данной статьи Тьюринга.

¹¹⁵ Семантическими можно считать те данные, которые играют роль предельного означаемого в некоторой системе координат — ясно, что это условное понятие, но столь же условно и понимание.

¹¹⁶ «Тот же самый мысленный эксперимент, — пишет Сёрл, — применим и к случаю с роботом» — Searle J. R. *Minds, brains, and programs* // *The Nature of Mind*. P. 514. Первоначально эта статья была опубликована в журнале *The Behavioral and Brain Sciences* (BBS), в 3-м номере за 1980 год.

¹¹⁷ Прекрасный обзор наиболее значимых возражений — Cole D. *The Chinese Room Argument: Entry in the Stanford Encyclopedia of Philosophy* 2004 <http://plato.stanford.edu/entries/chinese-room/>. См. также: *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence*, ed. by J. Preston and M. Bishop. N. Y., 2002; Mours J. *The Chinese Room Argument* // John Searle, ed. by B. Smith. Cambridge, 2003. P. 214-260. Среди критиков Китайской комнаты был и сам Сёрл. Речь идет не о том, что он отказался от этого аргумента, скорее он считает, что в нем он еще шел на поводу у когнитивистов, допуская объективную реальность информации, программ и компьютеров. В действительности, как доказывает он в ряде поздних работ, они существуют не сами по себе, а лишь при условии существования людей, т. е. наблюдателей, интерпретирующих те или иные материальные объекты в качестве программ, компьютеров или носителей информации — см., напр.: Searle J. R. *Mind*. P. 91-92. Однако эти доводы были давно опровергнуты — см.: Chalmers D. J. *The Conscious Mind*. P. 315-320, и я не буду принимать их во внимание (см. также: Edelman G., Tononi G. *A Universe of Consciousness: How Matter Becomes Imagination*. N. Y., 2000. P. 126-130). Я обобщил свои представления об аргументе Сёрла в статье «Кока-кола и секрет Китайской комнаты» — *Философия сознания: классика и современность*. М., 2007. С. 86-94.

¹¹⁸ Ограниченность подобных эмуляций была признана самим фактом развития коннекционистских моделей искусственного интеллекта, пришедших на смену символьным подходам, реализованным и в Китайской комнате. Но насколько существенны эти изменения, остается неясным.

¹¹⁹ Сходные возражения высказывались Сёрлу самыми разными авторами, но, как правило, в несколько смазанном контексте, а именно без акцентирования неэквивалентности Китайской комнаты и Китайского робота или же ограниченных эмулятивных ресурсов Комнаты. См., напр: Blackmore S. *Consciousness: An Introduction*. N. Y., 2004 (2003); P. 205-206; Harnad S. *Grounding symbols in the analog world with neural nets: A hybrid model* // *Think* 2 (1993). P. 12-20. Харнад, к примеру, признает, что Китайская комната может пройти тест Тьюринга безо всякого понимания китайского языка сидящим в ней человеком, но при этом утверждает, что Робот может пройти Тотальный тест Тьюинга (подразумевающий не только

языковую, но сенсорно-моторную компетентность) только при условии, что человек в нем действительно ощущает мир. Однако, чтобы опровергнуть этот вывод Харнада, не обязательно прибегать к экзотическим мысленным экспериментам (см.: Bringsjord S., Noel R. Real robots and the missing thought-experiment in the Chinese Room dialectic // Views into the Chinese Room. P. 144-166). В самом деле, если Китайская комната проходит тест Тьюринга, т. е. если возможно написать такую программу, которая позволит давать внешне осмысленные ответы на любые вопросы, касающиеся в том числе и обстановки этой комнаты, то вполне можно написать и такую программу, которая будет давать внешне осмысленные команды роботу в зависимости от его окружения, заранее прописанного в программе. Человек, управляющий таким чудесным роботом, может и не смотреть в его телекамеры, т. е. «видеть» его окружение, вовсе и не видя его.

¹²⁰ Searle J. R. Minds, brains, and programs // The Nature of Mind. P. 519: «Какими бы еще свойствами ни обладала интенциональность, она есть биологический феномен, и она должна так же каузально зависеть от специфической биохимии ее источников, как зависят от нее лактация, фотосинтез и другие биологические феномены». Интенциональность, напомним, по Сёрлу, необходимо связана с сознанием. О ее биохимических причинах см. также: *ibid.* P. 516.

¹²¹ См., напр.: Searle J. R. The Mystery of Consciousness. P. 202-203.

¹²² См.: Searle J.R. Mind. P. 114, 146-147.

¹²³ Тем не менее у этой идеи Сёрла до сих пор есть сторонники - см., напр.: Feinberg T. E. Altered Egos. P. 151.

¹²⁴ Ср.: Pinker S. How the Mind Works. P. 95; Rosenberg G. H. Re thinking nature: a hard problem within the hard problem // Explaining Consciousness — The "Hard Problem", ed. by J. Shear. Cambridge MA, 1997. P. 296-297.

Глава 3

ДЕННЕТ: А ВЫ ЗНАЕТЕ, ЧТО МЫ ЗОМБИ?

Дэниел Деннет совершенно не похож на Сёрла. Сёрл, к примеру, никогда не разместил бы на своем веб-сайте фотографию робота, как это сделал Деннет. Сёрл, скорее, кабинетный философ¹, Деннет ценит эксперименты. Сёрл известен в академических кругах, Деннет — большая медиа-звезда. Сёрл прозрачен, Деннет окутывает читателей туманными метафорами. Он гурман, он любит изысканные вина — Сёрлу же нравится холодное пиво. Несочетаемы у них даже хобби: Деннет бороздит океаны, а Сёрл летает на лыжах со снежных вершин. Хотя есть и совпадения. Так, оба автора уже обросли большой исследовательской литературой: о Деннете, к примеру, опубликовано не менее десятка специальных монографий и сборников², что же касается Сёрла, то один только Аргумент китайской комнаты вызвал целое море откликов. Есть у них и биографические пересечения. В частности, свою «книжную» философскую карьеру Деннет начал в том же году, что и Сёрл, — с работы «Контент и сознание» (1969). Их объединяет и ученический опыт в Оксфордском университете. Деннет, правда, на десять лет моложе Сёрла (он родился в 1942 г.), и его философские штудии в Оксфорде проходили в другой обстановке.

Сёрл, напомним, попал под влияние Остина и не впечатлился Райлом. Деннет, напротив, прошел райловскую школу философского бихевиоризма, к которой он был морально подготовлен в Гарварде, где преподавал лидер бихевиористской психологии Скиннер и ангажированный ею же знаменитый Уиллард Куайн — еще один выходец из Венского кружка, ученик Р. Карнапа, бросивший вызов учителю и всей аналитической философии критикой правомерности различения аналитических и синтетических предложений. Впрочем, «философ философов», как иногда называют Куайна³, трактовал проблему сознание—тело не так, как Райл, так что неправильно думать, что оксфордские учителя лишь повторяли Деннету гарвардские уроки⁴. В отличие от Куайна, полагавшего, что утверждения о ментальных состояниях могут быть просто проигнорированы⁵, Райл, как мы знаем, был уверен в возможности утилизировать их. Размышляя о собственной теории сознания, Деннет говорил, что она являет собой продукт скрещивания идей Куайна и Райла. И хотя это действительно так, следует учитывать, что куайновских элементов в ней

меньше. Квазикуайновским, помимо сциентистской установки, т. е. взгляда на науку как на «онтологического арбитра»⁶, у Деннета можно признать лишь тезис, что термины ментального словаря не допускают прямого перевода в слова о конкретных физических состояниях мозга⁷. Ментальные состояния напоминают, скорее, абстрактные центры гравитации или векторы сил. Но хотя сам Деннет придавал большое значение этой детали, оно не очевидно для тех, кто знакомится с его системой, особенно если при анализе опираются на его ранние или, наоборот, некоторые недавние работы⁸. Зато влияние Райла на Деннета нельзя не заметить. Хотя, разумеется, влияние это выражено не в том, что Деннет повторяет Райла — он, скорее, развивает его. Но основа у него райловская. И при таком положении вещей нам не обойтись без разъяснения позиции Райла.

Речь идет о позиции, занятой Райлом в уже упоминавшейся книге «Понятие сознания», которая вышла в свет в 1949 г. В предисловии к ее переизданию в 2000 г. Деннет заметил, что она относится к числу тех произведений, о которых все знают и говорят, но которые мало кто читает⁹. В самом деле: любому специалисту по аналитической философии известно, что в своей работе Райл сражается с декартовским «привидением в машине» и заявляет, что ментальные состояния можно истолковать как поведенческие диспозиции. Эта картина в общем верна, но она не позволяет в полной мере прочувствовать радикализм райловского проекта. А ведь это и правда был очень радикальный проект: чуть ли не каждая глава «Понятия сознания» начинается с обещаний ниспровергнуть давние традиции европейской философии в трактовке тех или иных психических способностей — воли, воображения, интеллекта и т. п.¹⁰

Более того: изгоняя ментальное «привидение» из машины человеческого тела, Райл вовсе не имеет в виду сказать, что он оспаривает субстанциальную теорию души, или духа как нематериального носителя ментальных состояний, и хочет заменить ее учением о том, что эти состояния при определенных условиях порождаются мозгом. Его задача гораздо более претенциозна: он пытается убедить читателей, что субъективных ментальных состояний вовсе не существует! Поначалу трудно даже поверить, что именно в этом его план, но это так.

Что хочет сказать Райл? Что означает несуществование, к примеру, ментальных образов?¹¹ Воображая зеленую корову, я ведь осознаю ее как некую воображаемую данность. Как можно отрицать это? Неужели можно настаивать, что ее не существует? Райл уверен: можно. Нет никакой воображаемой данности зеленой коровы, а есть определенный тип поведе-

ния, фантазирование. Но хотя со второй частью утверждения можно согласиться — в конце концов, человек, погруженный в свои фантазии, действительно ведет себя иначе, нежели, скажем, человек, разглядывающий реальную корову, — означает ли это, что фантазия — это только тип поведения? Ведь нам кажется, что мы непосредственно осознаем воображаемую корову как некую частную данность, доступную только нам самим. И это ключевой момент для понимания позиции Райла. Он вовсе не отрицает видимости наличия частных образов и других ментальных состояний. Нам, да и самому Райлу, признается он, кажется, что они существуют. Но это именно кажимость, иллюзия, не более того.

Некоторые философы, в частности тот же Серл, сталкиваясь с подобными рода утверждениями, сразу пытаются поймать их авторов на непоследовательности. В сфере ментального, утверждают они, не существует различия между видимостью и самой вещью. Если нам кажется, что у нас есть ментальные образы, то это равносильно признанию, что они действительно существуют — хотя бы в качестве кажимостей, и соответственно, нельзя утверждать, что их нет.

Это возражение выглядит убедительным, но все не так прозрачно. Дело в том, что сторонникам учения об иллюзорности ментальных состояний есть что ответить. Первая стратегия ответа состоит в том, чтобы подвергнуть сомнению «миф о данности»¹². Кажимость — вовсе не данность, а сложный конструкт, в котором имеются многочисленные теоретические привнесения. А если убрать их, ничего не останется. Ведь данности как таковой вообще не существует. Однако этот путь, на мой взгляд, лишен перспектив: отрицая правомерность понятия данности, мы неизбежно приходим к отрицанию возможности истинных высказываний, т. е. к самопротиворечию¹³.

Другой путь состоит в следующем: можно сказать, что, рассуждая об иллюзорности ментальных состояний, мы, не отрицая их наличия в качестве определенного рода данностей, имеем в виду их нестойкость и эфемерность, означающую, что при устранении причин, порождающих видимость их существования, они могут быть полностью элиминированы из жизни людей при сохранении адекватности их поведения. Это решение опробовала С. Блэкмор¹⁴. Конечно, оно предполагает, что видимость существования сознания порождается причинами, не имеющими существенного значения для человеческой жизни. В противном случае об иллюзии было бы говорить труднее и аргумент феноменалистов стал бы более опасным. Ведь они не отрицают, что определенные воздействия на мозг могут привести к выключению сознания, но

это еще не означает, что сознание иллюзорно, так как бессознательный мозг может и не породить поведение, которое продуцирует сознающий мозг. Вот если бы можно было исключить сознание, т. е. элиминировать ментальные состояния, без изменения возможностей мозга как механизма генерации поведения, то такие состояния по праву можно было бы считать иллюзорными. Но подобное исключение возможно лишь при условии, что приватные ментальные состояния не являются составными частями процесса формирования поведенческих реакций.

Понимание этого обстоятельства позволяет определить оптимальную стратегию критики позиции Райла и его сторонников (к которым принадлежит и Блэкмор, пусть и опосредованно, через Деннета), провозглашающих иллюзорную природу субъективных состояний. Критика эта не может ограничиваться «сёрловским» указанием на то, что в сфере ментального кажимость и реальность совпадают. Для придания действенности этому указанию оно должно быть дополнено демонстрацией того, что эти данности, взятые как таковые, в их реальном или кажущемся отличии от соответствующих им физических состояний (это не должно предрешаться), играют существенную роль в детерминации человеческого поведения, роль, которую нельзя описать на языке физических параметров мозга. Если бы удалось показать, что исключение приватных состояний при сохранении физического агрегата тела должно приводить к существенному изменению поведения, то говорить об их иллюзорности стало бы невозможным. В конце концов, реальным мы считаем такое явление, которое способно влиять на физические процессы: воображаемым молотком нельзя забить физические гвозди (хотя сам фантазийный образ молотка, конечно же, может быть реальным, но воздействующим уже не на гвозди, а на поведенческие процессы).

Итак, райловский тезис о несуществовании ментальных состояний, об их лишь кажущейся реальности, не может быть опровергнут простым указанием на его контринтуитивность, так как речь идет не о допущении или отрицании той или иной данности, а об интерпретации ее онтологического статуса, а здесь нет очевидности. Здравый смысл и впрямь утверждает, что внутренние ментальные состояния реальны, но здравый смысл говорил нам и о вращении Солнца вокруг Земли¹⁵. Не случайно сторонники Райла иногда сравнивают свой проект деконструкции ментальных состояний с научными революциями, подобными той, что осуществил Коперник. Укрепить позицию здравого смысла и прямо опровергнуть райловцев можно, подчеркнув еще раз, доказав каузальную

действенность качественных состояний, рассматриваемых как нейтральная данность. Возможности такого доказательства будут обсуждаться в пятой главе этой книги. Сейчас же обратим внимание на то, что райловцы, разумеется, тоже должны приводить какие-то аргументы: согласимся ведь, что противоречие здравому смыслу — это еще не критерий истинности. Причем им надо не только доказывать иллюзорную природу сознания, но и выявлять причины возникновения этой иллюзии.

Самый простой путь разделаться с этими вопросами — негласно объединить две задачи и решить вторую (как это делали радикальные элиминативисты¹⁶) указанием на то, что иллюзия существования сознания как совокупности разного рода квалиа порождается самим фактом наличия и применения ментального словаря. Однако из этого следует, что для избавления от иллюзии сознания достаточно прекратить использовать такие слова, как «боль», «образ» и т. п. Но нелепо полагать, что отказ от употребления слова «боль» избавит нас от страданий. Подобные объяснения нельзя принимать всерьез¹⁷.

Райл действовал более тонко. Он вовсе не призывал отказываться от ментального словаря. Он анализировал «логическое поведение»¹⁸ ментальных терминов (в том числе и отсюда именование его философии «логическим бихевиоризмом», хотя в этот термин обычно вкладывается другой смысл¹⁹) и пытался показать, что оно не соответствует тому, которое можно было бы ожидать от них, если бы референты этих терминов подпадали под онтологическую категорию, аналогичную категории физических явлений. Их подлинным (а не кажущимся) референтом оказываются не вещи, а диспозиции, как, к примеру, в случае «настроений», или сложные поведенческие схемы, как в случае с «образами», некоторыми «чувствами», «рассуждениями» и т. п. Если Райл прав, то и впрямь можно утверждать, что, говоря о ментальном как сфере приватной данности, мы совершаем категориальную ошибку, порождающую ментальное «привидение», — и остается лишь объяснить, почему мы ее совершаем.

Райл выдвигает на первый план исторические причины. В XVII веке многие ученые поверили в то, что физический мир допускает исчерпывающее толкование с механистических позиций и, соответственно, что природа — это всего лишь громадная машина. Машиной некоторые мыслители, вроде Гоббса, признали и человека. Другие, более религиозно настроенные ученые, прежде всего Декарт, не согласились с этим и попытались вывести ядро человеческого Я за пределы механистического мира. Это ядро, душа, была объявлена особой непространственной вещью, носителем нематериаль-

ных качеств и способностей. Так и возник, будучи компромиссом механистического естествознания и ветхой религии, миф о «привидении в машине». Райл, кстати, считает, что этот миф был шагом вперед, сыгравшим прогрессивную роль²⁰. Он стимулировал правильную мысль, что человеческая личность не может быть представлена в качестве совокупности грубых физических свойств. Человека нельзя объяснять ньютоновской механикой. Он - нечто большее, чем простой механизм, машина. Другое дело, что это нечто большее вовсе не есть какой-то духовный довесок, как думал Декарт, а набор поведенческих диспозиций, природу которых в состоянии определить лишь современная биология.

Таким образом, вырисовывается следующая картина. Миф о «привидении» как продукт XVII века был, по Райлу, неизбежным результатом совместного действия трех факторов — нового естествознания, старой религии и отсутствия внятных биологических теорий, появившихся лишь в XIX веке благодаря Ч. Дарвину и Г. Менделю²¹.

Но из этих объяснений остается неясным, почему даже после появления нужных биологических теорий философам так трудно избавиться от «привидения в машине»? Почему от него трудно избавиться даже самому Райлу? Он не скрывает, что его книга есть некий опыт аутотерапии. И терапия нужна, судя по всему, не только философам — ведь, как следует из пояснений Райла, этот миф существует и на обыденном уровне. На этот уровень он был внедрен религией, считает Райл. Но только ли дело во внешних влияниях? Размышления о мучительном расставании Райла с «привидением» наводили его учеников на мысль, что стойкая вера людей в существование особой ментальной реальности имеет и другие причины. Самая простая из них могла бы состоять в том, что такая реальность действительно существует. Возможно, райловский анализ ментальных состояний имеет лишь ограниченную истинность, думали они. Возможно, некоторые из этих состояний, вроде желаний или настроений, действительно могут быть истолкованы как поведенческие диспозиции, но это еще не значит, что в таком ключе могут быть интерпретированы все ментальные феномены. В конце концов, сам Райл признавался, что не удовлетворен своими объяснениями некоторых из них, а именно ощущений и чувств, или переживаний²². Из подобных сомнений выросла, как мы знаем, знаменитая теория тождества.

Теоретики тождества — Плейс, Фейгл и Смарт — признавали существование субъективных ментальных состояний²³, но заявляли, что они могут быть по факту отождествлены с состояниями мозга. Однако эта уступка мифу о привидении

не пришлась по вкусу Райлу и его ортодоксальным сторонникам, прежде всего Деннету. Неудивительно, что он начинает свою книгу «Контент и сознание», написанную под непосредственным руководством Райла, с критики теории тождества. Эта экстравагантная теория, утверждает Деннет, может быть принята только как крайнее средство в случае провала всех попыток поведенческого объяснения ментальных состояний²⁴.

Вместе с тем Деннет считает, что анализ этих состояний — в том виде, в каком он был предложен Райлом, — недостаточен. Его требуется углубить и провести на ином уровне, не ограничиваясь рассмотрением обыденного употребления терминов ментального словаря. Ведь подобное рассмотрение заставляло Райла утверждать, что ментальные термины отсылают к внешним поведенческим схемам — дальше этих схем обыденное употребление и не уходит. Однако такой подход фактически толкал Райла на классические бихевиористские позиции, подразумевающие под поведением совокупность внешних реакций на определенные стимулы, — позиции, бесперспективность которых в 60-е годы XX века стала уже очевидной.

И дело было не только в том, что сам Райл плохо справился с бихевиоризацией некоторых ментальных состояний. Причины гораздо шире. В то время многие уже осознали, что любое определение ментального состояния через предрасположенность к поведенческим реакциям должно дополняться отсылкой к другим ментальным состояниям. К примеру, желание прочитать книгу может быть определено как предрасположенность взять эту книгу с полки, раскрыть ее и т. п., но здесь обязательно надо оговариваться, что эта предрасположенность предполагает определенные убеждения относительно этой книги (которые, в свою очередь, предполагают какие-то другие убеждения), а также то, что она может быть реализована в поведении лишь при условии отсутствия других, более сильных желаний.

Одним словом, большинству авторов, пишущих на эти темы, и в том числе Деннету, к тому времени стало ясно, что стремление переопределить ментальные состояния через внешние поведенческие эквиваленты приводит лишь к тому, что на месте одних внутренних состояний, от которых удается избавиться, словно драконьи головы, вырастают другие внутренние состояния, включаемые в итоговую дефиницию. Выходило, что ментальные состояния невозможно дефилировать без учета внутренних факторов, игнорируемых традиционным бихевиоризмом. И чтобы спасти стратегию Райла, Деннету надо было решиться на учет этих факторов, но отрицать их

ментальную природу, попытавшись определить их в физических терминах так называемых «центральных состояний», в терминах процессов, происходящих в мозге, — в этом плане проект Деннета был близок плану действий Д. Армстронга²⁵. Поскольку речь шла не о деталях, а о концептуальном моделировании, Деннет не видел нужды углубляться в нейро-биологические частности и собирался ограничиться более абстрактным уровнем разработок новомодной в то время когнитивной науки. Но он не забывал и о намеках Райла на важность биологии, и прежде всего теории эволюции в философии сознания. Мозг — биологический объект, и именно теория эволюции может ответить на многие вопросы, связанные с его устройством и функциями.

Теперь мы уже в состоянии зафиксировать стартовые позиции деннетовской философии сознания. С одной стороны, как ученик Райла Деннет оборонял его идеи от критики его неортодоксальных последователей, прежде всего теоретиков тождества, вновь запустивших «привидение» в машину только затем, чтобы тут же попытаться крайне загадочным способом отделаться от него. С другой стороны, чтобы защитить Райла, он должен был пойти гораздо дальше своего учителя, сместиться с «персонального» уровня анализа ментальных состояний (уровня обыденного языка или «народной психологии») на «субперсональный» уровень концептуального моделирования работы мозга и сменить вальяжный костюм лингвистического философа на спецовку когнитивного ученого и неodarвиниста.

Только что нарисованная картина, впрочем, нуждается в немедленной корректировке, которая позволит сделать ее более реалистичной. Во-первых, неверно утверждать, что Деннет был непримиримым оппонентом создателей теории тождества и целиком отвергал ее. Более того, в интервью, опубликованном в 2006 г. в журнале *Minds and Machines*²⁶, он даже заявил, что в свое время развивал идеи Смарта, ведущего теоретика тождества. Этот тезис, конечно, не вполне историчен (хотя мы видели, что в самой теории тождества изначально существовала элиминативистская райловская линия). В действительности в «Контенте и сознании» мы не находим апологии Смарта. Но правда в том, что уже здесь Деннет отмечал, что знаменитую фразу Смарта, что состояниям мозга тождественны не сами ментальные образы, а «опыт их наличия», можно истолковать в его, деннетовском ключе²⁷. Во-вторых, специфику своей позиции он пояснял, опираясь в предварительном плане именно на аргументы лингвистической философии обыденного языка. А специфика эта в том, что термины ментального словаря являются, по Деннету, «не-

референциальными», т. е. не отсылающими к особым ментальным сущностям, допущение которых возродило бы «привидение».

Как же можно показать неререференциальную природу ментальных терминов? Разве такое допущение не контринтуитивно? Разве не означает оно отрицание реального существования боли и радости, желаний и убеждений как специфических частных феноменов? Деннет согласен с тем, что он идет против интуиций. Он согласен, что людям кажется, что они обладают частными состояниями. Здесь, как мы видим, он в полной мере следует Райлу. Но даже анализ некоторых нементальных слов демонстрирует, что обыденные интуиции не всегда гарантируют истинность основанных на них утверждений. В качестве примера Деннет рассматривает слово «голос»²⁸. На первый взгляд очевидно, что на вопрос «существуют ли голоса?» — разумеется, не те, что слышались Орлеанской деве, — надо отвечать положительно. Но что такое голос? — спрашивает Деннет. Четкое определение трудно отыскать, так как самое правдоподобное из них: «членораздельные звуки, издаваемые вибрацией голосовых связок» (я обобщаю целый ряд деннетовских вариантов) легко отбрасывается, учитывая то, что голоса могут передаваться, к примеру, по радио. Подобная размытость значения, впрочем, характерна для большинства слов, и для Деннета это лишь повод продемонстрировать особенности его подхода. Из неясности значения слова «голос» не надо делать вывод, что голоса — некие неуловимые, хотя и в полной мере реальные сущности. Скорее, надо пойти против интуиции, которая, как видно, может и сбивать с толку, и заключить, что у этого слова вообще нет прямого референта и что оно обладает значением лишь в качестве составной части более широких идиоматических контекстов, которым можно подыскать физические соответствия.

То же самое могло бы иметь место и с терминами ментального словаря. Если бы удалось показать, что эти слова всегда включены в определенного рода контексты вроде смартовского «опыта наличия образов» и что эти контексты допускают переопределение в терминах физических состояний мозга, то тезис о несуществовании референтов ментальных слов, т. е. о несуществовании особых ментальных сущностей сам получил бы право на существование, несмотря на свою контринтуитивность. Впрочем, Деннет не забегает вперед. Он предлагает начать с нейтрального рассмотрения ментальных терминов и посмотреть, можно ли в принципе надеяться на переопределение их контекстов на языке физической науки. Более того, по сути он предлагает исходить из установок сторонников тези-са о реальности ментальных состояний и их резком отличии

от физических явлений. Говоря на эти темы, такие феноменологически настроенные мыслители с подачи Brentano и Husserl обычно апеллировали к понятию интенциональности как сущностному свойству ментального. Деннет принимает вызов, соглашается с этим утверждением и приступает к обсуждению природы интенционального отношения.

В слове «интенциональное» в английском, да и в русском языке слышится как «намеренное», так и «нацеленное на что-то». Близость этих смысловых оттенков, впрочем, очевидна, хотя второй имеет более широкий охват: нацеленность можно понимать как направленность на предмет, как наличие какого-то содержания. Именно этот момент обыгрывает Деннет. Ментальные состояния, признает он, — это такие состояния, в которых есть содержание, «контент». К примеру, желание — это не просто желание, а желание чего-то, надежда — это надежда на что-то, восприятие - восприятие чего-то, и т. п.

«Чтойность» во всех этих формулах как раз и выражает предметную, контентную сторону соответствующих состояний. Констатация контентной природы ментальных состояний позволяет Деннету конкретизировать его неорайловскую программу физикализации последних. Чтобы показать возможность перевода ментального языка на язык физических процессов, надо по меньшей мере убедиться в том, что эти процессы могут быть истолкованы таким образом, что можно будет по праву приписывать им какое-то содержание, значение.

Разумеется, приписать какое-то значение можно любым физическим событиям — средневековые авторы любили развлекаться в подобном ключе. Но есть ли такие ситуации, когда, отвлекаясь от всякой произвольности, можно сказать, что некие физические процессы действительно значат то-то и то-то, несут такой-то и такой-то контент? Если такие ситуации могут быть обнаружены, то якобы непроходимая пропасть между интенциональными ментальными и неинтенциональными физическими состояниями будет преодолена — именно потому, что важнейшее свойство ментального, интенциональность, контентность, окажется совместимым с функционированием сугубо физических систем. Это и означало бы принципиальную возможность перевода контекстов ментальных терминов на физикалистский язык и отрицания наличия у этих ментальных терминов особых нефизических референтов, т. е. попросту означало бы демонстрацию возможной иллюзорности нашего сознания.

Трудно сказать, какое развитие получила бы у Деннета идея отыскания физических эквивалентов интенциональных

идиом, если бы перед глазами у него не стояли новые компьютерные технологии, интерес к которым Деннет ощутил уже в начале 60-х. Компьютеры в те времена еще никто не наделял внутренним миром ментальных состояний (ситуация стала менее определенной после работ Чалмерса), никто не заселял их райловскими привидениями. Тем не менее в ряде случаев было очевидно, что мы имеем полное право говорить об их функционировании на интенциональном языке. Любимый, как он сам признает, пример Деннета — шахматные программы. Играя с виртуальным оппонентом, практически невозможно не приписывать ему интенциональных состояний. Если хотите победить, то приходится допускать, что такой противник тоже нацелен на выигрыш, что он собирается атаковать вашего короля, защищаться и т. п. Лишь приписывая шахматному компьютеру желания, страхи и намерения, можно эффективно предсказывать его действия. Данная стратегия, таким образом, полностью оправданна. Но при этом все или почти все согласны, что компьютер — просто железяка, лишенная сознания.

Это означает, что даже такое приписывание ментальных состояний, которое лишено произвольности, не исключает возможности объяснения функционирования подобных интенциональных систем на физикалистском языке, языке полупроводников и электрических схем. Для Деннета этот примечательный факт — залог возможной продуктивности исследовательских усилий, позволяющих физикализировать и человеческую интенциональность и показать отсутствие различий между «внутренней интенциональностью» и «как бы интенциональностью», — различий, на которых, как мы помним, энергично настаивал Сёрл. Но в только что упомянутом факте, по Деннету, можно найти и другие подсказки. Если шахматная программа в чем-то похожа на человеческое сознание, то, взглянув на дело с другой стороны, вполне можно предположить, что и сознание, возможно, в чем-то похоже на программу, а мозг — на компьютер.

Важность данной аналогии в том, что у программы есть программист. Именно он ответствен за возможность интенци-ональной интерпретации физических процессов, протекающих в системном блоке компьютера. Именно его действия делают такие процессы носителями какого-то контента и цели. И если сознание аналогично программе, то программист должен быть и у него. И программист этот в общем известен: это дарвиновский естественный отбор. В самом деле, естественный отбор способствует формированию поведенческих схем, реализация которых благоприятствует выживанию и продолжению потомства организмов в данных условиях. Иными

словами, естественный отбор закрепляет целесообразное, «разумное» поведение, которое обычно и описывается с помощью интенциональных идиом, на языке ментальных, сознательных состояний.

Круг тем самым замкнулся. Еще раз пробежимся по нему. Чтобы не допускать особый класс ментальных сущностей и не попадать на маятник взаимоисключающих и одинаково, по Деннету, абсурдных альтернатив дуализма, материализма теории тождества и идеализма, мы должны допустить возможность перевода ментальных контекстов на язык физических феноменов. Возможность такого перевода предполагает, что некоторым физическим состояниям можно приписывать контент или значение. Но утверждать, что физический феномен имеет какое-то значение для некоего организма, можно, лишь если этот организм адекватно, «разумно» реагирует на него. Формирование таких реакций объясняется действием естественного отбора, а сами эти реакции закрепляются на уровне нейронных связей в мозге. Схемы таких нейронных связей могут быть истолкованы как своего рода программы, управляющие поведением организма.

Такова база деннетовской теории сознания — в том виде, в каком она была представлена в первом разделе «Контента и сознания». Но базой, разумеется, нельзя ограничиться, на ней надо что-то строить. В самом деле, пока Деннет показал лишь, что тезис о несуществовании особого онтологического класса ментальных состояний мог бы претендовать на истинность, несмотря на его контринтуитивность. Но от возможности хорошо бы перейти к действительности, а для этого надо показать, что квазисамодостаточные ментальные сущности и в самом деле сводимы к функциональным схемам мозга.

Деннет решает эту задачу во втором разделе «Контента и сознания», озаглавленном «Сознание». Здесь Деннет начинает с различения двух видов «осознания», а именно: (1) того, которое проявляется в способности к адекватным невербальным поведенческим реакциям, и (2) того, собственно человеческого сознания, индикатором которого служат вербальные отчеты; затем последовательно рассматривает различные ментальные состояния и действия — ментальные образы, намерения, мышление и т. п., — и показывает, как можно было бы попытаться отыскать их физикалистские первоосновы.

Учение о сознании составляет, таким образом, вторую часть системы Деннета — первой же, напомним, является учение об интенциональности, или контенте. Вот как об этом говорит сам Деннет: «Моя основополагающая стратегия всегда оставалась неизменной: сначала разработать объяснение

контента как чего-то независимого и более фундаментального, чем сознание... а затем построить на этой базе объяснение сознания. Сначала контент, затем сознание»²⁹. В последующие после выхода «Контента и сознания» годы его система выдержала еще два «издания», обрастая все новыми деталями и уточнениями. Первое из них пришлось на 1978 год, когда Деннет опубликовал «Мозговые штурмы», композиционно организованные по модели «Контента и сознания». Дальнейшее развитие его теории сделало ее настолько масштабной, что ему пришлось отказаться от попыток упаковать ее в одной монографии, и он разбил ее третью инсталляцию на два этапа, реализованных в «Интенциональной установке» (1987) и «Объясненном сознании» (1991). Тезисы «Объясненного сознания» получили развитие в сборнике «Сладкие грезы» (2005), а учение об интенциональности было доработано Деннетом в «Опасной идее Дарвина» (1995) и «Видах психики» (1996). Попробуем суммировать основные нововведения Деннета в поздних работах.

В эссе «Автопортрет», опубликованном в 1994 г. в «Справочнике по философии сознания» и переизданном в авторском сборнике «Дети мозга» (1998), Деннет отмечает, что вторая часть его системы, теория сознания, подвергалась с 1969 г. более существенным переделкам, чем учение об интенциональности³⁰. По сути, он дает понять, что изложение теории сознания в «Контенте и сознании» в ряде аспектов было попросту неадекватным. Начать с того, что, как утверждает сам Деннет, в 1969 г. он преувеличивал «роль языка в конституировании частных, но существенных аспектов человеческого сознания»³¹. Особенно это заметно по деннетовской трактовке отношения языка и ментальных образов. В «Контенте и сознании» он пытался показать, что эти якобы «фотографические» внутренние картинки в действительности являются лишь дескрипциями, речевыми формулами, которые проще редуцировать к механизмам мозга. Обосновывал это Деннет тем, что, к примеру, воображая некоторые вещи, мы четко фиксируем их воображаемые свойства только тогда, когда даем воображению соответствующую речевую инструкцию. Скажем, представьте зебру. А теперь подсчитайте ее полосы. Как ни странно, сделать это будет непросто. Деннет утверждал, что затруднения неслучайны: без четких речевых указаний о полосках и подобных им качествах их число остается неопределенным. Уже один этот факт разрушает сходство ментальных образов с изображениями. В общем, заключал он, нет независимых от языковых инструкций ментальных фото — уверенность в их бытии есть просто иллюзия. Позже Деннет утверждал, что зашел в этой трактовке «слишком уж далеко»³². По-

хоже, что он прав. Подсчитать полоски на зебре нельзя и тогда, когда мы видим ее периферическим зрением - но из этого вовсе не следует, что периферическое зрение конструируется речью, скорее, ее участие в нем минимизировано.

Однако помимо частных недоработок в «Контенте и сознании», по мнению его автора, была допущена и более масштабная ошибка: там предполагалось, что в мозге существует некая невидимая, хотя и вполне реальная линия, разделяющая бессознательные и сознательные процессы³³. Деннет называл ее «линией осознания» (awareness line) и говорил о принципиальной возможности отыскания физической «границы» речевого центра, пересекая которую тот или иной контент становится доступен для выражения и осознания³⁴.

Эта концепция, содержащая рудименты «картезианского материализма», как сказал бы поздний Деннет, была позже заменена им «теорией множественных набросков», впервые детально представленной в «Объясненном сознании». Но, прежде чем обратиться к анализу этой работы Деннета, поговорим о той части его теории, которая составляет фундамент или предпосылку учения о сознании и которая не пересматривалась, а лишь дорабатывалась им после 1969 г.

Концепция интенциональности — речь идет о ней — действительно претерпела определенные изменения после «Контента и сознания». Первое из них пришлось на 1971 год, когда Деннет опубликовал статью «Интенциональные системы», вошедшую позже в сборник «Мозговые штурмы». Развивая здесь тезис «Контента и сознания» о приписывании убеждений, желаний и других интенциональных состояний некоторым материальным системам в целях эффективного предсказания их поведения, он называет такой подход «ин-тенциональной установкой» и отличает ее от «физической» и «дизайнерской» установок³⁵. Физическая установка срабатывает тогда, когда для предсказания определенных событий достаточно воспользоваться общими представлениями о законах природы, как, к примеру, в случае предсказания падения камня при выпускиании его из рук³⁶. Дизайнерская установка применима к артефактам, функции и механизмы которых хорошо известны. Зная, для чего нужен будильник и что он такое, я могу предсказать, что он зазвонит в такое-то время.

Рассуждая о будильнике и подобных ему устройствах, нам совершенно незачем приписывать им желания и убеждения. Ситуация меняется, когда мы видим, что какая-то материальная система демонстрирует целесообразное «поведение» во внешней среде. Представим, к примеру, механические устройства на колесах, спроектированные так, что они избегают столкновения со стенами комнаты и друг с другом (я видел та-

ких танцующих роботов на художественной инсталляции, приуроченной к Туссанской конференции 2006 г.³⁷). Размышляя о них, хочется сказать, что они стремятся избежать столкновений, т. е. использовать определенную интенциональную идиому и принять по отношению к ним интенциональную установку.

Заметим, в случае с такими роботами и подобными им простыми устройствами, вроде термостатов, искушению говорить о них в интенциональных терминах легко противостоять. Однако ему почти невозможно противиться, когда принципиальное устройство аналогичных материальных систем неизвестно нам или когда это устройство очень сложно, как у компьютеров. Приписывая им интенциональные состояния, мы, стало быть, словно заполняем имеющиеся у нас пустоты в знании об их физической организации³⁸, обретая вместе с тем хорошие шансы на предсказание их поведения.

Именно это обстоятельство привлекает Деннета. Жизнь людей — сплошное прогнозирование, и для него хороши любые средства, лишь бы они были эффективными. Но уроки термостатов и танцующих роботов учат нас, что приписывание материальным системам интенциональных состояний нельзя абсолютизировать. И то, что сегодня рассматривается как интенци-ональная система, завтра может быть лишено примысленного внутреннего мира. Одна установка может сменяться другой³⁹. Поняв устройство живых существ, мы сможем предсказывать их поведение на основе дизайнерской установки, а в предельном случае — и на основе физической установки.

Ключ к принципиальному устройству живых существ дает знание законов эволюции: формирование их целесообразного, «разумного» поведения осуществляется в процессе естественного отбора. Но отбор целесообразных поведенческих схем не происходит одновременно, это длительный процесс, и важно уяснить его этапы. Деннет решает эту задачу в «Опасной идее Дарвина» и «Видах психики». Он выделяет четыре этапа эволюции поведения. На первом существа жестко генетически соотнесены со своим наличным окружением. Такие организмы Деннет называет «дарвинианскими созданиями?». На втором этапе возникают «скиннерианские создания», способные варьировать поведение в зависимости от положительных или отрицательных подкреплений своих конкретных действий. Для третьего этапа характерно появление «поппериан-ских созданий», умеющих «проигрывать» будущие действия в своей внутренней информационной среде еще до их реального совершения и позволяющих «гипотезам умирать вместо них». На этой стадии возникают высшие животные. Но она еще не предел. Информационная среда, в которой проходит

проигрывание их гипотез, содержит помимо врожденных, генетически детерминированных схем в основном данные их индивидуального опыта. Между тем все мы прекрасно понимаем, что эту среду можно обогатить и коллективным опытом современников и предков, который называется культурой. Но внедрение в мозги индивидов культуры предполагает существование некоего медиума, способного транслировать этот коллективный опыт. Речь, разумеется, идет о языке. Именно «языковой инстинкт», присущий людям, позволяет им передавать или декодировать различные культурные технологии и достигать качественно нового уровня насыщения информационной среды их нейронных компьютеров. Поэтому люди заслуживают особого места в плане эволюции. Деннет называет их «грегориjsкими созданиями» (по имени психолога Р. Грегори)⁴⁰.

Итак, если мы хотим объяснить человеческое сознание, мы должны, считает Деннет, принимать во внимание культуру. Но как работать с ней? Долгое время в европейской философии доминировала квазигегелевская традиция гипос-тазирования культуры, рассмотрения ее как некоей самостоятельной сущности, «мирового духа», возвышающегося над индивидами. Лишь во второй половине XX века эта те-оцентрическая в своих истоках идея, преломлявшаяся в самых разных концепциях, в том числе в марксизме, стала уходить со сцены. Возможно, решающий удар нанес по ней британский биолог Ричард Доукинс. Произошло это как бы ненароком. В 1976 г. он опубликовал книгу «Эгоистичный ген», разошедшуюся с того времени тиражом более миллиона экземпляров. Там Доукинс доказывал, что субъектом естественного отбора являются не индивиды или виды, как считали многие, а гены или комплексы генов. Т. е. в процессе естественного отбора отбираются такие гены, которые обладают способностью к порождению максимального числа копий. Понимание этого обстоятельства позволяет метафорически изобразить всю биологическую эволюцию как борьбу эгоистичных генов. Виды при таком подходе оказываются абстракциями, а индивиды — переносчиками генов, подчас жертвующими собой ради блага боссов.

Рассуждая на эти темы в общебиологическом плане и лишь к 11 главе добравшись до человека, Доукинс без колебаний признает, что людей нельзя объяснить исключительно с биологических позиций, поскольку они существуют не только в мире природы, но и в мире культуры. Эта порядком затасканная мысль обретает, однако, у него совершенно новое звучание, когда он отмечает, что допущение им автономности культуры вовсе не означает, что сама культура не мо-

жет быть объяснена с помощью эволюционной теории. Ведь теория эволюции не обязана ограничиваться биологическими объектами.

Иными словами, помимо генов могут быть и другие репликаторы, сущности, порождающие собственные копии. И по мнению Доукинса, в культуре есть аналоги генов. Он называет их «мемами». Собственно, культура из них и состоит. В ме-мах нет ничего таинственного. Их примерами могут быть «мелодии, модные словечки и выражения, способы варки похлебки и сооружения арок»⁴¹, в общем, любые «идеи», способные к реплицированию, пересаживанию от одного человека к другому. Мемы могут рассматриваться как хозяева (точнее, эгоистичные паразиты) наших мозгов, они борются за существование в условиях ограниченных ресурсов (возможности человеческих мозгов небеспредельны), имеют разную способность к выживанию и могут мутировать при копировании. Эволюция культур может быть, по Доукинсу, объяснена из этих посылок.

Все эти мысли лишь мимоходом высказаны в «Эгоистичном гене», и сам Доукинс утверждал, что все значение своей гипотезы он осознал по итогам ее детальной проработки Деннетом, после которой многие всерьез заговорили о создании новой науки — «меметики».

Но хотя Деннет действительно подхватил идею мемов и придал ей новый импульс, он осторожно относится к перспективе создания меметического учения о культуре как двойника дарвиновской теории эволюции. Истоки его сомнений не имеют для нас большого значения. Главное, что при всех этих колебаниях он уверен, что концепция мемов как элементов культуры может работать независимо от того, как далеко простирается их аналогия с генами. И наиболее эффективной областью ее приложения является теория сознания.

Чтобы понять, почему это так, обратим внимание на устройство мемов. Каждый мем есть что-то вроде инструкции или программы. Именно «программная», «алгоритмическая» природа мемов позволяет им пересаживаться из одного мозга в другой. Изучение, имитация — вот основные способы освоения культуры, но выучить и симитировать можно лишь то, что может быть представлено в виде определенной, т. е. отличной от других, последовательности шагов.

Ясно, что представление о культуре как программном пуле, некоем бульоне алгоритмических инструкций и технологий, инсталлирующихся подобно компьютерному софту в человеческие мозги, великолепно сочетается с деннетовским проектом науки о сознании. Если «обычные» животные по большей части довольствуются генетически вшитыми про-

граммами, дополняемыми индивидуальным опытом, то человеческий мозг каким-то образом сочетает природный багаж поведенческих программ с тем, что подгружается в него извне. Разгадка человеческого сознания — в уяснении этих механизмов.

Здесь мы уже выходим за пределы первой части деннетовской системы, общего учения об интенциональности. Обсуждение механизмов программного наполнения человеческого мозга и порождаемых им продуктов осуществляется Деннетом в учении о сознании. В наиболее развернутом виде это учение, напомним, изложено в «Объясненном сознании», и сейчас нам предстоит более подробно поговорить об этой книге. В конце предыдущей главы я назвал ее одним из бесспорных шедевров современной мысли. В ней есть все, что требуется для выдающегося философского произведения: масштабность замысла, смелость решений, оригинальность и остроумие.

Тем не менее сама моя оценка, конечно, не бесспорна. Далеко не все философы присоединились бы к ней. Показательна, к примеру, позиция Д. Чалмерса. В начале 2005 г. он пригласил посетителей своего блога «Фрагменты сознания» поучаствовать в составлении списка лучших книг по философии сознания, вышедших в двадцатом столетии. Предваряя дискуссию, он поделился собственными предпочтениями. Три книги «выбрались сами»: «Понятие сознания» Райла, «Материалистическая теория сознания» Армстронга и «Сознание и его место в природе» Брода (ее мы еще коснемся), но дальше у него возникли трудности. Перебирая разных авторов, в том числе и Деннета, Чалмерс дает понять, что считает лучшими его книгами «Контент и сознание», «Мозговые штурмы» и «Интенциональную установку», а вот «Объясненное сознание» он называет «важной книгой в плане ее влияния в данной области, но, совершенно очевидно, не столь успешной в философском плане, как предыдущие»⁴². Эта фраза Чалмерса может показаться странной, но я могу понять его. Чалмерс считает, что в «Объясненном сознании» Деннет не выполнил свое обещание — не объяснил сознание, подменив решение «трудной проблемы» обсуждением «легких проблем сознания». И Чалмерс не одинок в этом суждении. Подобные вещи можно услышать от Н. Блока, М. Маккинси, да и других философов. Насколько справедливы эти суждения, нам сейчас и предстоит выяснить.

Начнем наш обзор с композиции «Объясненного сознания». После предисловия — где Деннет сообщает, что «после тридцати лет раздумий, разговоров и создания текстов» о тайне сознания он наконец может предложить «набросок

решения»⁴³, — и «Прелюдии», где он, в частности, уточняет, что собирается «объяснить различные феномены, составляющие то, что мы называем сознанием, — показывая, как все они оказываются физическими эффектами мозговых процессов, как эволюционировали эти процессы и как они порождают иллюзии о своих силах и свойствах»⁴⁴, Деннет переходит к последовательному разрешению обрисованных выше задач.

В первой части книги — «Проблемы и методы» — Деннет предлагает осуществить путешествие в «феноменологический сад», в котором обитают те самые ментальные сущности, совокупность которых мы называем сознанием, - чувства, эмоции, образы и т. п. Рассматривая все эти хорошо знакомые каждому феномены, он постоянно обращает внимание на то, что многие из них оказываются не тем, чем они обычно считаются. К примеру, стандартное описание феноменального зрительного поля предполагает равномерную «окрашенность» всех его частей. Однако самый простой эксперимент — попытка определить цвет вещи, находящейся на его периферии, — показывает, что это иллюзия. Тем не менее Деннет настаивает, чтобы исходным пунктом объяснения сознания была именно та картина феноменологического сада, которая рисуется обыденными его реконструкциями, отчетами людей о своих состояниях⁴⁵. С таких отчетов, принимаемых на веру, и должен начинать философ — Деннет называет подобную практику «гетерофеноменологическим методом».

Этот термин породил множество недоумений и толкований. В самом деле, в чем, собственно, отличие гетерофеноменологического метода от «аутофеноменологического», когда философ говорит о своих собственных ментальных состояниях? По мнению Деннета, гетерофеноменологический метод позволяет рассуждать о сознании с позиции стороннего наблюдателя, с точки зрения третьего лица, и тем самым претендовать на научность своих построений. Задача гетерофеноменолога — объяснить, почему люди говорят о сознании то, что они говорят. И кажется, что рассуждения о сознании — как особый вид поведения — действительно можно объяснять с позиции третьего лица. Проблема, однако, в том, что, прежде чем начать подобные объяснения, надо понять смысл таких рассуждений, а понять их можно, лишь приписав людям внутренние ментальные состояния. Но приписывать другим людям какие-то внутренние состояния можно, лишь перенося на них собственный феноменальный опыт. Поэтому разница между гетерофеноменологией и аутофеноменологией скорее номинальна, чем реальна. Максимум, о чем можно говорить, так это о различии акцентов. Рассуждая о сознании

другого человека, легче согласиться с тем, что его внутренний мир — иллюзия. Вспомним «шахматную» аналогию Деннета: приписывая ментальные состояния виртуальному оппоненту, мы в любой момент готовы отыграть назад и сказать: на самом деле их нет. Но, быть может, можно отыгрывать назад и с реальными оппонентами, людьми? Занимая аутофено-менологические позиции, отыграть труднее. Кажется, что, приравнивая себя к компьютеру, мы лишаем себя чего-то очень существенного, нашего внутреннего мира. Деннет с этим не спорит: действительно, так нам кажется, — но кажимость еще не истина.

В общем, гетерофеноменологический подход, в отличие от аутофеноменологического метода, позволяет, по Деннету, не так болезненно расставаться с иллюзиями сознания. Поэтому он его и практикует⁴⁶. Впрочем, до расставания с иллюзиями еще далеко. Во второй части «Объясненного сознания», озаглавленной «Эмпирическая теория сознания», Деннет презентует свой главный объяснительный инструмент — «теорию множественных набросков», одним из положений которой является тезис о населяющих мозги мемах.

Третья часть книги, названная Деннетом «Философские проблемы сознания», демонстрирует, каким образом указанная теория позволяет демистифицировать сущности, населяющие «феноменологический сад». Иными словами, используя теорию множественных набросков, Деннет показывает здесь, чем в действительности являются ментальные образы, квалиа⁴⁷ и, наконец, сами наши личности. Он утверждает, что «дисквалификация» квалиа и других ментальных сущностей происходит именно при опоре на упомянутую теорию, хотя на деле в соответствующих разделах базой является скорее общая элиминативистская идеология этой теории. По идее, в третьей части Деннет должен объяснять и то, как и почему порождается иллюзия существования особого ментального мира.

В действительности краткий ответ на вопрос «почему?» - и состоящий в том, что эта иллюзия, возможно, выгодна в эволюционном плане⁴⁸, — дан лишь в приложении. Так что я не случайно использовал выражение «по идее» и т. д. Ведь, обращаясь к самому тексту «Объясненного сознания», мы видим, что реальная композиция этой книги расходится с образом, который был нарисован выше со слов Деннета. Очевидно, к примеру, что изгнание «привидения» ведется уже во второй части: теория множественных набросков изначально подается Деннетом на контрасте с менталистской доктриной «Картезианского театра», который она и призвана демонтировать. А первый критический заход сделан Деннетом в пер-

вой части. Реальная структура его аргументации видится мне примерно так. Прежде всего Деннет нападает на классический картезианский дуализм, который не устраивает его потому, что не в состоянии объяснить возможность влияния духовной субстанции на мозг и, соответственно, ведет к пораженческим настроениям. Иными словами, хотя Деннет признает, что не может опровергнуть дуализм, он считает, что, поскольку дуалист фактически расписывается в неспособности объяснить сознание — в том числе и потому, считает он, что объяснение возможно лишь через сведение того, что объясняется, к чему-то другому, — принятие этой доктрины непродуктивно⁴⁹.

Вслед за этим Деннет показывает, что, хотя большинство современных авторов разделяет его антидуалистические настроения, далеко не все из них осознают, что отказ от дуализма должен вести к радикальной перестройке учения о сознании. В самом деле, многие из тех, кто отвергает дуализм, а значит признает зависимость сознания от мозга, считают возможным сохранить рудимент духовной субстанции, говоря о некоем Центре ментальных процессов, в котором происходит порождение смыслов, отдача распоряжений и, самое главное, презентация ментальных состояний, словно на сцене. Это место, обиталище Я, Деннет и называет Картезианским театром.

Упомянутая позиция, идентифицируемая Деннетом как «картезианский материализм», была бы возможна лишь при наличии в мозге некой «точечной» структуры, в которую стекалась бы вся афферентная, т. е. идущая с периферии тела и самого мозга, информация, и которая вместе с тем оказывалась бы точкой отсчета для нисходящих, эфферентных поведенческих и иных управляющих команд. Деннет, однако, убежден, что такое представление не подтверждается нейронаукой: это, по его мнению, «очевидно плохая идея»⁵⁰. Сразу следует отметить, что это суждение не выглядит бесспорным: в конце концов, нашумевший проект Крика и Коха, приложивших немало сил для отыскания нейронных коррелятов сознания, по крайней мере не противоречит идее о том, что структуры такого рода все-таки могли бы реально существовать в мозге⁵¹. Но если допустить, что Деннет прав, что это будет значить? Ответ Деннета достоин тщательного обдумывания, хотя кажется, что он является простым выводом из ранее озвученных положений; если в мозге отсутствует Театр, в нем не может быть никакой Сцены и Зрителей.

Раз в мозге нет зрителей или Зрителя, наблюдающего за Презентацией ментальных состояний, то что же в нем есть? Есть множество нейронных процессов. Они протекают одно-

временно: мозг как биологический орган, обрабатывающий информацию, т. е. входные данные, поступающие на органы чувств, имеет параллельную архитектуру. Многие из этих процессов, схемы которых возникли под давлением естественного отбора, на генетическом уровне соотнесены с соответствующими поведенческими реакциями. Однако для сложных существ, наподобие человека, характерна высокая пластичность действий. Это означает, что поведенческие схемы не обязательно должны быть прошиты в hardware мозга: они могут возникать либо вследствие аутостимуляции индивида, либо загружаться в мозг извне, из культуры, которую, как мы уже знаем, Деннет считает громадным программным пулом, до предела насыщенным мемами.

Дальше начинается самая интересная часть теории сознания Деннета. Он заявляет, что мемы, инсталлированные в мозг, формируют там особую «виртуальную машину» и что эта машина создает иллюзию потока сознания. Но здесь, конечно, нужны уточнения. Многое проговаривает сам Деннет. Прежде всего, отмечает он, «из интроспекции» известно, что наше сознание организовано как «джойсовский» поток, т. е. его состояния выстроены в некую линию, а не даны одновременно. Это значит, что порождающий его «компьютер» имеет последовательную, «фон-неймановскую» архитектуру⁵². Но мозг — параллельный компьютер. Выходит, сознание является продуктом «вычислительной машины» более высокого уровня, установленной на параллельную архитектуру мозга. Такое устройство, стало быть, — временная структура, «сделанная скорее из правил, чем из проводов»⁵³. Подобные структуры, образцом которых на обычных компьютерах являются операционные системы, такие, как Windows или текстовые процессоры (Word и т. п.), Деннет и называет «виртуальными машинами». Если машина, генерирующая поток сознания, действительно устроена так, то она должна быть привнесена в мозг извне. А извне — значит из культуры, океана мемов. Вполне логично поэтому выглядит утверждение Деннета, что «само человеческое сознание есть громадный комплекс мемов (или, точнее, мемо-эффектов в мозге) и оно лучше всего может быть понято в качестве действия "фон-неймановской" виртуальной машины»⁵⁴.

Но почему виртуальная машина, порождающая сознание, имеет последовательную архитектуру? Этот вопрос⁵⁵ не очень удобен для Деннета. Дело в том, что сама виртуальная машина была постулирована им для объяснения линейности состояний сознания. Т. е. он исходил из того, что мозг как параллельный компьютер не может сгенерировать подобный феномен. Однако, поскольку виртуальные машины не обязаны

иметь последовательную архитектуру, ему не удастся уйти от вопроса, почему мемы, загруженные в мозг, не создают в нем параллельную виртуальную машину. Исходя из (1) «культурного» характера виртуальной машины, из того, что (2) мемы, составляющие культуру, передаются в основном через язык, а также из того, что (3) первоначальным стимулом к образованию «софтверных» связей в мозге могли, по Деннету, послужить — неожиданно продуктивные — диалоги человека с самим собой, можно предположить, что последовательный характер виртуальной машины он будет объяснять, отсылая к особенностям человеческой речи, тем более что ее трудно помыслить иначе, как используя образ линейного развертывания. Неудивительно, что некоторые интерпретаторы Деннета ухватились за этот вариант объяснения⁵⁶. Однако если Деннет и склонялся к нему в начале 90-х гг., через десять лет его позиция изменилась. Когда я прямо спросил его, связана ли последовательная, «фон-неймановская» архитектура виртуальной машины с линейной структурой речи, вместо «да» он ответил нечто другое (письмо от 4 августа 2004 г.): «Главная причина того, что это "фон-неймановская" машина заключается в том, что она зависит от "дорогостоящих" массовых или монополизирующих эффектов ("эффектов славы") в мозге, а они являются конкурентными в использовании доступных ресурсов и тем самым создают нечто вроде бутылочного горлышка».

Таким образом, получается следующая картина. Некоторые процессы в мозге забирают так много ресурсов, что они не могут протекать одновременно, и виртуальная машина превращает их в последовательные. Вдумавшись в эту фразу, мы заметим, что особой необходимости в виртуальной машине и формирующих ее мемах в таком случае не остается: по определению упомянутые процессы не могут быть одновременными. Данное обстоятельство сразу проблематизирует учение Деннета о культурной обусловленности человеческого сознания. И это означает, что в «Объясненном сознании» Деннет едва ли мог придерживаться той концепции, которую он упомянул в процитированном письме. Изменение его позиции могло быть связано с осмыслением новых экспериментальных данных⁵⁷. К примеру, из теории, изложенной в «Объясненном сознании», следует, что лишь человеческое сознание — именно в силу его культурной, меметической основы — организовано как поток. Однако хорошо экспериментально подтвержденный эффект «бинокулярного соперничества» (*binocular rivalry*) — когда два глаза видят разные картинки, но при этом происходит не наложение этих картинок, а их последовательное восприятие, — иллюстрирующий

линейную структуру сознания, наблюдается не только у человека, но и, скажем, у обезьян⁵⁸, у которых явно отсутствуют деннетовские виртуальные машины.

В свете этого и других фактов — сам Деннет дает понять, что на изменение его позиции повлияли, скорее, не конкретные эксперименты, а весь массив свежих эмпирических данных⁵⁹, — неудивительно, что в сборнике «Сладкие грезы» Деннет называет тезис о виртуальной машине в мозге как конститутивной части сознания «смелым утверждением», которое может и не выдержать проверку опытом⁶⁰. При этом он отделяет данный тезис от другой, логически не зависимой от него, более важной и, как ему кажется, надежной гипотезы, объясняющей суть сознания. Речь идет о концепции «фантазийного эха». Она тесно связана с более ранней гипотезой «славы в мозге», которая эволюционировала из теории «множественных набросков», подробно изложенной в «Объясненном сознании». Учитывая некоторые высказывания Деннета, можно было бы подумать, что в данной ситуации точнее говорить не о смене концепций, а, скорее, лишь о замене метафор⁶¹. Однако, как мы только что видели, в поздних работах Деннет стал четче дифференцировать различные компоненты своей теории сознания и придавать им разное значение.

Присмотримся к тому блоку концепции Деннета, который кажется ему наиболее существенным. Сначала коснемся его на материале «Сладких грез», а затем вернемся к «Объясненному сознанию». Итак, «фантазийное эхо». Это словосочетание обозначает удерживание в воображении тех или иных событий, имевших место в прошлом, или, иными словами, эпизодическую память. Нейронные процессы, фундирующие это «эхо» (и в свою очередь, возможно, фундированные культурным влиянием, осторожно добавляет Деннет), таким образом, как бы не затихают в мозге, сохраняют там свое «влияние». Иными словами, они словно бы известны всему мозгу, «прославлены» в нем. Но именно о таких, «прославленных» в мозге процессах мы говорим, что осознаем их, заявляет Деннет. Это утверждение не является чем-то необычным в наши дни. Напротив, оно, скорее, выражает, как отмечает сам Деннет⁶², некий консенсус, наметившийся среди когнитивных ученых в последние десятилетия. Самый известный лейбл для подобной трактовки сознания придумал, впрочем, не Деннет, а Б. Баарс, о котором уже упоминалось ранее. Я имею в виду его модель «глобального рабочего пространства». Еще в 80-е годы он пришел к выводу⁶³, что сознательными процессами в мозге становятся тогда, когда их продукты словно бы транслируются там, становятся доступными самым разным его подсистемам, т. е. попадают в глобальное рабочее пространство.

В дальнейшем многие когнитивисты и нейрочеловеков поддержали Баарса.

Собственно, деннетовская теория «славы в мозге», детальным образом прописанная в сборнике «Сладкие грезы»⁶⁴, мало в чем отстывает от модели Баарса⁶⁵. Впрочем, на предыдущем этапе различий было больше. Хотя в «Объясненном сознании» Деннет тоже пытался вписать свою теорию в рамки модели глобального рабочего пространства⁶⁶, он насыщал ее множеством оригинальных деталей. Он не только акцентировал внимание на необходимости особой виртуальной машины, устанавливающей порядок исполнения глобальных нейронных программ (напомню, что в поздних текстах необходимость такой культурно обусловленной машины подвергается сомнению: упорядочение могло бы происходить и за счет генетически детерминированных устройств), но и специфически понимал сам механизм их исполнения. Особенность этого механизма иллюстрирует понятие «наброска», которое он использовал в тот период. Мозг можно представить как сообщество специалистов, или «демонов», каждый из которых обрабатывает свой узкий участок. Итогом такой работы и оказываются «наброски» — некие протосуждения, фиксирующие то или иное содержание, различенное этими «демонами».

Судьба таких «набросков» может быть разной: одни «играют скоротечную роль в модуляции текущей активности, другие, быстро сменяя друг друга, выдвигаются на новые функциональные роли»⁶⁷. В числе таких ролей может быть и формирование индивидуумом вербальных отчетов о своих внутренних состояниях. «Наброски», о которых этот субъект может отчитаться, обычно и наделяются статусом осознанных, но в «Объясненном сознании» Деннет настаивает, что такая привилегия не всегда оправдана. В действительности, утверждал он — явно в духе известного куайновского тезиса о неопределенности перевода, — у нас нет «единого, определенного потока сознания»⁶⁸. И нет его потому, что в мозге нет «Центрального офиса, Картезианского театра, "где все сходится вместе"»⁶⁹. Есть множество набросков, и какой из них, к примеру, назвать осознаваемым раньше, какой — позже, если речь идет о малых промежутках времени, это во многом дело произвола⁷⁰. Вообще при чтении «Объясненного сознания» возникает ощущение, что осознанность того или иного состояния, по Деннету, не является четко фиксированным свойством⁷¹. Если взять «пробу» потока сознания одним способом, субъект скажет, что он осознает что-то одно, если другим — совсем другое. Кроме того, нарративные наброски, продуцируемые в мозге, находятся в процессе постоянного редактирования, переплетаясь с другими набросками. Мы все время переделываем свое сознание.

Деннет не отрицает последний тезис и в более поздних работах. В этом смысле его версия теории глобального рабочего пространства как «славы в мозге» все же отличается от ее «консенсусного» варианта, предполагающего, что для сознания достаточно трансляции контентов в этом пространстве⁷². И она допускает интерпретацию в духе концепции множественных набросков - Деннет не случайно вспоминает слова Энди Уорхола: «В будущем каждый сможет стать знаменитым на 15 минут». По Деннету, эта фраза абсурдна: слава только тогда слава, когда она продолжительна⁷³.

Но как понять, продолжительной ли будет слава, если она только начинается? Что, если эту начинающуюся славу затмит какое-то другое событие? Она уже не будет славой. Так и с сознанием. Процесс, вышедший в глобальное рабочее пространство, может быть вскоре вытеснен каким-то другим процессом. Можно ли в таком случае будет говорить, что тот первый процесс был осознанным? Не обязательно. А это значит, что понятие осознанного состояния может быть относительным и в теории «славы в мозге»⁷⁴. Но из этого не следует, что между теорией славы в мозге и теорией множественных набросков нет различий. Ведь согласно первой из них — или по крайней мере согласно одной из ее правдоподобных интерпретаций — множественные наброски, могущие быть истолкованными в качестве осознанных, сосуществуют, тогда как вторая ограничивается допущением возможности их быстрой смены в глобальном рабочем пространстве. Можно, правда, вспомнить, что и теория множественных набросков признает факт последовательной смены сознательных состояний, факт «спрямления» этих набросков виртуальной машиной. Но суть в том, что, согласно этой теории, спрямление является лишь иллюзорным. Теория славы в мозге считает его реальным⁷⁵.

Таким образом, в работах конца 90-х — начала 2000-х гг. Деннет пришел к чему-то вроде синтеза своих более ранних концепций. С одной стороны, он фактически отказался от радикальной идеи иллюзорности потока сознания, представленной в «Объясненном сознании», и тем самым сделал шаг в сторону «Контента и сознания», где он допускал наличие реальной «линии», разделяющей осознанные и неосознанные представления. С другой стороны, теперь он понимает эту линию — и само глобальное рабочее пространство, — скорее, не анатомически, а динамически. По словам С. Хамероффа, на конференции в Лас-Вегасе в июне 2007 г. Деннет утверждал, что осознанным может стать чуть ли не любой процесс в мозге - при условии его большей интенсивности в сравнении с другими⁷⁶. При такой трактовке неудивительно, что одним

из условий осознанности он считает длительное проигрывание в мозге разного рода процессов.

На различия исходной теории множественных набросков и теории славы в мозге обратила внимание в своей рецензии на «Сладкие грезы» уже известная нам сторонница Деннета — С. Блэкмор⁷⁷. Она резонно сочла, что Деннет отошел от своих оригинальных плюралистических тезисов в сторону более стандартной концепции, и задумала остроумный эксперимент (результаты которого были опубликованы в 2006 г.), который мог бы вернуть Деннета к его изначальным идеям. Идея эксперимента состояла в следующем: если теория множественных набросков в ее исходном виде верна, т. е. если в мозге нет единого потока сознания, то при опрашивании субъекта, который делает два дела сразу (к примеру, ведет машину — реально или виртуально — и слушает музыку), относительно того, что он осознавал непосредственно перед контрольным сигналом, результаты будут различаться в зависимости от того, какого рода сигнал им получен: если это звуковой сигнал, то он, скорее, скажет, что осознавал музыкальную композицию, если зрительный — то дорогу перед собой. Однако проведенные замеры показали, что корреляция такого рода не обнаруживается⁷⁸. И это нанесло еще один удар по теории множественных набросков, хотя сама Блэкмор призывает не делать окончательных выводов, что неудивительно: по сути, она атаковала саму себя, поскольку Деннет уже отошел от этой теории.

Мне, впрочем, не хотелось бы, чтобы у читателей создавалось впечатление, будто теория множественных набросков Деннета оказалась совершенно уж несостоятельной. Некоторые ее части до-прежнему выглядят весьма убедительно. К примеру, удачным представляются деннетовские объяснения механизмов порождения речи или алгоритмов принятия решений. Он пытался показать, что традиционные картины, предполагающие существование генерирующего осмысленную речь и выносящего решения Я, не соответствуют реальной феноменологии процессов подбора слов и воления и гораздо лучше объясняются с позиций теории множественных набросков. В самом деле, речь формируется, скорее, не сверху, из мифического Я, а снизу, из множества протонарративных конкурентных процессов — именно поэтому слова и фразы обычно как бы выныривают на нас, возникают из каких-то глубин. Так же и с принимаемыми решениями. Интересно выглядит и тезис Деннета о том, что единое Я, иллюзия которого существует у людей, есть лишь абстрактный «центр нарративной гравитации». В реальности, полагает он, в этой связи можно говорить только о системе, или «петле», монито-

ринга⁷⁹, с одной стороны, и динамических фокусах мозговой активности — с другой: если некий набросок выдвинулся на первый план (или оказался назначен на своего рода «капитанскую» должность виртуальной машиной⁸⁰), то именно он становится временным пристанищем нашей «самости». Но она быстро соскальзывает в другое место, и т. д.

В общем, трудно спорить с тем, что представление о едином Я как устойчивом центре осмысления и воления перегружено проблематичными теоретическими допущениями. Деннет мастерски снимает эти спекулятивные напластования. Вопрос лишь в том, не заходит ли он слишком далеко. Прежде всего неясно, как он может допускать существование в мозге системы мониторинга и при этом отрицать, что в нем имеется место, куда сходятся те информационные потоки, которые мы так или иначе осознаем. Возможная реплика такова: система мониторинга существует, однако она сугубо материальна, в ней нет ментального наблюдателя. Допустим, но откуда это известно? Деннет дает ответ и на этот вопрос, но что это за ответ? Вспомним канву его рассуждений: поскольку в мозге нет места, «где все сходится воедино», в нем нет Зрителя и Сцены, где разыгрываются ментальные Представления. А значит, не существует и самих ментальных представлений, квалиа, а есть только мозг как физический орган с мириадами параллельных процессов обработки информации. В частности, мозг обрабатывает информацию о собственной активности. Результаты такой обработки называют интроспективными отчетами. Специфика работы человеческого мозга такова, что в этих отчетах неизбежно появляются слова о потоке сознания, о Я, о состояниях, которые кажутся чем-то качественно отличным от физических процессов в мозге (так как система мониторинга не имеет доступа к низкоуровневым «демонам», генерирующим подобные состояния). Такие суждения понятны, но их не надо принимать за чистую монету, и уж тем более превращать их в какие-то вечные истины.

Вдумавшись в эти конструкции, мы увидим, что они содержат противоречие: система мониторинга, куда стекается вся осознаваемая информация, не включает в себя наблюдающее Я, так как Я не существует, а не существует оно потому, что оно могло бы существовать лишь при наличии в мозге места, куда стекается вся информация, а такого места не существует. Избежать противоречия можно попробовать, сказав, что система мониторинга не может играть роль стационарного Картезианского театра: это нечто подвижное. Но это не помогает, так как подвижным может быть и Я. Даже если в мозге нет фиксированного центра схождения информационных потоков, из этого еще не следует, что в нем нет динамического

центра, а значит, и динамического Зрителя. Такого Зрителя допускает и сам Деннет, но, обратим внимание, уже после того, как он сделал онтологические выводы о несуществовании качественных ментальных состояний. Т. е. он рассуждает так: нет фиксированного центра — нет Театра, нет Театра — нет Зрителя, нет Зрителя — нет ментальных Представлений, нет ментальных представлений или состояний — есть только физические процессы, причем некоторые из них могут осуществлять мониторинг за другими процессами, взявшими на время власть в мозге, и в результате этого мониторинга порождать иллюзию Зрителя (т. е. совокупность вербальных отчетов о «привилегированных» состояниях), а затем и теорию неизменного Зрителя.

Это настоящий паралогизм: из отсутствия фиксированного центра не следует отсутствие Зрителя — а лишь прикованного к одному месту Зрителя, напоминающего узника платоновской пещеры, и стационарной Сцены. Это дополнение существенно, так как отсутствие стационарной Сцены, в отличие от отсутствия Сцены как таковой, не допускает однозначной онтологической интерпретации. Если Сцены нет вообще, то, конечно, не существует и ментальных Представлений, квалиа, В противном же случае нет достаточных оснований опустошать мозг от качественных состояний и фактически объявлять людей «зомби», пусть и такими, которые даже и не подозревают об этом. А Деннет делает подобный шаг: «Возможны ли зомби? Они не только возможны» они действительны. Все мы зомби»⁸¹. В других местах он именует людей «зимбо» — это особая разновидность зомби с продвинутыми механизмами самомониторинга⁸².

Тема зомби так популярна, что нельзя на время не задержаться на ней. Иногда можно услышать мнение, что зомби в огромном количестве населили философские тексты благодаря усилиям Д. Чалмерса. Но хотя он и в самом деле приложил руку к раскрутке этих странных героев, само понятие зомби как лишённого сознания существа, являющегося, однако, точной физической и поведенческой репликой человека, обсуждается давно. Еще Декарт касался этой темы — разумеется, не употребляя слова «зомби». Однако он считал, что человеческое тело не способно само по себе, без участия души, продуцировать поведение, характерное для людей, в частности продуцировать речь. В XVIII веке ситуация изменилась, и, к примеру, Хр. Вольф в своей «Метафизике» доказывал, что, несмотря на внешнюю нелепость понятия бессознательных физических копий людей, они могли бы существовать. В XX веке эта тема была акцентирована в 1974 г. р. Кирком, который и ввел термин «зомби» в широкий философский обо-

рот⁸³. Д. Чалмерс считает, впрочем, что терминологическая новация Кирка оставалась практически незамеченной до появления его собственных работ в середине 90-х гг. Однако это мнение нельзя принять. В самом деле, достаточно взглянуть на обсуждение проблемы зомби в «Объясненном сознании» Деннета. Примечательно, что в этой книге, вышедшей, напомним, в 1991 г., ее автор говорит о термине «зомби» как о чем-то хорошо известном и даже заявляет об «общем согласии среди философов» в том, что «зомби являются или были бы такими людьми, которые демонстрируют совершенно естественное, сопровождаемое вниманием и речью, живое поведение, но при этом в действительности полностью лишены сознания, будучи чем-то вроде автоматов»⁸⁴.

Деннет, впрочем, вовсе не считает, что понятие зомби может принести реальную пользу в философских дискуссиях. Напротив, оно, скорее, запутывает их. Дело в том, полагает он, что это понятие основано на иллюзии в той мере, в какой оно контрастируется с понятием человека. В действительности между людьми и «философскими зомби», дефиниция которых была дана выше, вообще нет различий. Ведь сознания, которого якобы лишены зомби, просто не существует, а в том смысле, в каком оно существует, зомби обладают им. Именно поэтому при желании всех людей можно назвать зомби.

Итак, Деннет настаивал и настаивает на том, что между людьми и зомби не существует реального различия. Вера в такое различие, утверждал он в статье «Интуиция зомби» (2001), перепечатанной в сборнике «Сладкие грезы», основывается исключительно на традиции⁸⁵ и, скорее всего, будет устранена в течение ближайшего столетия. Впрочем, сама порочная традиция обособления сознания, как он полагает, имеет глубокие корни, и это означает, что полного ее исчезновения ожидать не приходится: «она не сохранится в своей нынешней, токсичной форме, но устоит в качестве менее опасной мутации, все еще психологически мощной, но утратившей влияние»⁸⁶. Подобно тому как люди научились противостоять видимости вращения Солнца вокруг Земли, они привыкнут и к мысли об иллюзорности существования сознания как особой реальности: «По-прежнему кажется, — будут, по мнению Деннета, изрекать они, — что эти механистические теории сознания оставляют что-то за бортом, но, разумеется, это иллюзия. В действительности они объясняют все, что требуется объяснить относительно сознания»⁸⁷.

Деннет упрекает «реакционных» философов, таких, как Чалмерс или Дж. Левин, отстаивающих «зомбизм», в нежелании слушать голос разума, предполагая, по-видимому, что его позиция имеет серьезные аргументативные основания.

Мы видели, однако, что некоторые его аргументы напоминают паралогизмы. Впрочем, его общая установка сохраняет хотя бы видимость правдоподобия: по крайней мере, мыслимо, что углубление нашего знания о реальном функционировании мозга приведет нас к ситуации, когда для предсказания поведения людей будет совсем не обязательно приписывать им ментальные состояния. Т. е. в обыденных ситуациях это по-прежнему будут делать, но с оговоркой, что это просто удобное допущение⁸⁸. В таком случае тезис о том, что другие люди — зомби, уже не покажется чем-то неординарным, как в наши дни нам не кажется неординарным тезис о «зомбий-ности» компьютеров или роботов.

Но видимость правдоподобия установки Деннета сохраняется, лишь пока мы говорим о других людях, — а как быть с самими собой? Приведет ли новое знание о функциональных схемах мозга к ситуации, когда мне не нужно будет приписывать сознание самому себе? Но приписываю ли я вообще себе сознание? Кажется, что для подобного приписывания у меня уже должен быть какой-то материал, те самые квалиа, которыми я наделяю других. А если они уже есть, то, значит, собственное сознание не приписывается, а изначально дается⁸⁹.

Однако, несмотря на то, что это возражение звучит убедительно, позицию Деннета можно попробовать защитить. Все дело в том, что осознанные состояния подозрительно тесно связаны с вербальными отчетами. Если я осознаю нечто, то могу сказать об этом. Данное обстоятельство позволяет предположить, что сами квалиа могут быть интерпретированы как вербальные конструкции. Т. е. вполне может оказаться, что мое представление о собственных квалиа все же опосредовано суждениями. Но если это так и если суждения, в свою очередь, могут быть истолкованы в терминах функциональных диспозиций — что вполне вероятно, — то картина сразу усложняется.

В самом деле, при таком истолковании проблема зомби выглядит следующим образом: при недостатке знаний о механизмах своего собственного функционирования мозг, точнее система мозга, ответственная за внутренний мониторинг, естественным образом продуцирует суждения о том, что интегральная информация, получаемая этой системой, лишена необходимой связи с низкоуровневыми процессами в мозге. Сама эта информация трактуется как нечто качественно отличное от генерирующих ее нейронных сетей. Человек в таком состоянии утверждает, что обладает неким особым внутренним миром, специфика которого не может быть зафиксирована с позиции третьего лица — возникает понятие квалиативного. Однако при расширении его знаний о собственном мозге он начинает понимать,

как именно низкоуровневые процессы в мозге связаны с его интегральными контентом, фиксируемыми системой мониторинга. И хотя это понимание или знание тоже представлено в виде «ментальных образов» и состояний, т. е. в качестве высокоуровневых данностей, — и таким образом полного устранения иллюзии качественное не происходит, — но теперь он может уверенно утверждать, что это именно иллюзия и что на самом деле так называемые квалиа вовсе не являются чем-то, что нельзя познать с позиции третьего лица, т. е. объективными научными методами.

Если это рассуждение корректно, то различие между людьми и зомби действительно утрачивается. Ведь зомби, будучи по определению бихевиориальными дубликатами обычных людей, с необходимостью будут генерировать у себя ту же самую иллюзию наличия у них качественных состояний. Точнее, они будут говорить, что у них есть ментальные состояния и что они кажутся им отличными от физических состояний их мозга. Но если наши качественные состояния действительно опосредованы суждениями о них и сконструированы этими суждениями — а в конечном счете бихевиориальными диспозициями, — то мы ничем не будем отличаться от них.

Все, таким образом, упирается в вопрос о том, можно ли согласиться с данным тезисом — об опосредованности нашего знания о собственных ментальных состояниях суждениями на этот счет. Нельзя отрицать, что в некоторых случаях такая опосредованность имеет место, и Деннет показывает это. Речь идет об уже упоминавшейся теоретической нагруженности ряда представлений о наших внутренних состояниях. Но из того, что зачастую мы видим наши ментальные состояния сквозь «теоретические» очки, не следует, что при их удалении вообще не остается никаких ментальных качественных данностей или состояний. Более того, в ряде случаев мы вообще не знаем, как вербализовать состояния подобного рода — если сталкиваемся, к примеру, с совершенно новыми переживаниями. Кроме того, кажется — и это, на мой взгляд, есть решающий аргумент, — что даже при совершенно точном понимании, какого рода процессы в мозге связаны, к примеру, с наличием визуальных образов, представления об этих процессах не будут включать в себе элементы самих этих образов. Представление нейронной основы ментального образа, к примеру, голубой коровы не будет содержать в себе голубизны — разумеется, не является голубой и сама эта нейронная основа. И это значит, что всегда можно говорить еще о чем-то, кроме как о состоянии мозга, а именно о том, «каково это», по уже упоминавшейся формуле Т. Нагеля, находиться в таком состоянии.

Ситуацию можно проиллюстрировать мысленным экспериментом «Нейроученая Мэри», предложенным в 1982 г. австралийским философом Фрэнком Джексоном в блистательной статье «Эпифеноменальные квалиа»⁹⁰. Эксперимент этот обсуждался почти так же живо, как Аргумент китайской комнаты Сёрла, причем в итоге этих обсуждений Джексон отказался от своих изначальных выводов, а именно от тезиса о ложности физикализма⁹¹. Суть мысленного эксперимента Джексона в следующем: представим милашку Мэри, выросшую в черно-белом окружении и занимающуюся изучением нейронных механизмов зрения. Мэри — девушка из будущего, и она живет во времена, когда нейронаука полностью прояснила эти механизмы. Иными словами, Мэри знает все о зрении с его, так сказать, физической стороны. Но она никогда не видела ни голубых, ни зеленых, ни красных предметов. И вот в один прекрасный день она выходит из своей черно-белой лаборатории и видит мир. Вопрос: узнает ли она что-то новое о зрении, столкнувшись с окрашенными предметами? Джексон утверждал, что да, она узнает, каково это видеть цвета, и делал отсюда вывод, что в мире, помимо физических фактов, известных Мэри и ранее, есть и нефизические факты, примером которых, в частности, могут быть качественные данные о цвете, полученные ею после выхода из черно-белой комнаты.

Впоследствии Джексон согласился со своими многочисленными критиками, что сделанный им онтологический вывод является слишком сильным и что он непосредственно не вытекает из принятых им посылок. Мэри обретает что-то новое, с этим трудно спорить, но это может быть не новое знание, а новая способность⁹², или новые понятия о старых фактах⁹³. Любопытно, однако, что даже эти смягченные выводы не устраивали Деннета. Он всегда утверждал, что Мэри вообще не узнает ничего нового по выходе из комнаты. Деннет цеплялся за тезис, что Мэри знает все физические факты о зрении (он был нужен Джексону для чистоты вывода), и доказывал, что при таком фантастическом в наши дни предположении мы уже не сможем доверять нашему обыденному интуитивному представлению о том, что она расширит свое знание после выхода из комнаты⁹⁴.

Но почему Деннет вообще занимал столь бескомпромиссную позицию? Очевидно, потому, что ее малейшее смягчение приносило бы качественный элемент в теорию восприятия, от которого он хотел избавиться. Да вот только можно ли последовательно отстаивать ее?

Предположим, что Мэри и правда знает о цветном зрении все, в том смысле, что ей доподлинно известно, что при воспри-

ятии объектов, называемых, к примеру, зелеными, в мозге в определенной последовательности активируются такие-то группы нейронов. Но отсюда, разумеется, не следует, что это знание с необходимостью активирует данные группы нейронов в ее собственном мозге. Подобная активация, конечно, возможна, но считать ее необходимой — значит допускать неразрывную связь, с одной стороны, нейронов, реализующих это знание, с другой — отвечающих за цвет⁹⁵. Но трудно понять, на каком основании мы должны допускать эту связь.

Итак, хотя Мэри в состоянии предвосхитить, что именно будет происходить в ее мозге после того, как она увидит, к примеру, лужайку, покрытую свежей травой, это предвосхищение может, тем не менее, не сопровождаться появлением у нее образа зеленого цвета — именно потому, что, по предположению, он возникает только в случае активации определенных групп нейронов, тех самых, которые могут остаться у нее дезактивированными. Поэтому, хотя, глядя на свой мозг со стороны после выхода из комнаты, она не найдет ничего нового, она действительно может заметить что-то новое в ином аспекте. Если бы между людьми и зомби не было различий, само наличие этого качественного аспекта, который, как мы видим, Деннет не имел права игнорировать, было бы невозможным. А поскольку он есть у людей — хотя бы у гипотетической Мэри и у автора этих строк (за Деннета ручаться нельзя)⁹⁶, — люди все-таки должны быть отличны от зомби, при том что из построений Деннета неясно, почему люди обладают этой качественной добавкой и как она возникает.

В «Сладких грезах» сам Деннет — явно переосмысляющий в последние годы свою позицию и все более и более энергично заявляющий, что ему самому кажется, будто эта добавка имеет место, — фактически признает, что его теория не дает ответа на этот вопрос⁹⁷. При этом он оговаривается, что считает его несущественным. Однако именно этот вопрос выражает «трудную проблему сознания», о чем, кстати, упоминает и сам Деннет⁹⁸. Одним словом, выходит, что, как и Сёрл, Деннет, собиравшийся объяснить сознание, в итоге не делает этого и оставляет «трудную проблему» без решения⁹⁹. Точнее, так — Деннет предлагает нам альтернативу: либо «трудная проблема» (1) нерешаема, но не имеет особого значения, либо мы все же зомби и тогда она (2) вообще не существует¹⁰⁰, а автор, прославивший ее, воспользовался приемом одного фокусника, трюк которого не могли разгадать даже профессионалы — именно потому, что они считали, что это один «трудный» фокус, тогда как в действительности речь шла о наборе разных «легких» трюков¹⁰¹. Однако даже эта установка все равно не снимает «трудной проблемы»: Деннет не забывает

подчеркивать, что, хотя, с его точки зрения, он смог продвинуться в направлении разложения «трудной проблемы сознания» на множество «легких» проблем, на этом пути он не достиг окончательных результатов¹⁰². И теперь нам не остается ничего другого, как обратиться к текстам того философа, который, собственно, и изобрел «трудную проблему», — в надежде, что в них мы найдем реальные попытки ее решения. Речь, разумеется, идет о Дэвиде Чалмерсе. Анализ его взглядов и посвящена следующая глава.

Примечания

¹ В более поздних работах Сёрл, впрочем, все чаще ссылается на экспериментальные данные и все более определенно заявляет, что «полагает не всегда возможным резко разграничивать концептуальные и эмпирические вопросы» — Searle. J. R. *Freedom and Neurobiology*, P. 30-31.

² На русском языке о Деннете есть большая монография Н.С. Юлиной «Головоломки проблемы сознания» (2004), на мой взгляд точно передающая дух деннетовского философствования. Лучшее англоязычное введение в философию Деннета — Zawidzki T. *Dennett*. Glasgow, 2007. Автор этой книги справедливо называет Деннета системосоздателем. Более аналитичная монография — Elton M. *Daniel Dennett: Reconciling Science and Our Self-Conception*. Cambridge, 2003.

³ Этот титул Куайн заслужил потому, что он писал явно не для широкой публики. И это была его осознанная позиция — см.: Quine W. V. *Theories and Things*. Cambridge MA, 1981. P. 190-193.

⁴ После защиты диссертации в 1965 г. в Оксфорде Деннет вернулся в США. С 1971 г. он работает в Тафтском университете, расположенном в пригородах Бостона.

⁵ См.: Quine W. V. *The Ways of Paradox*. N. Y., 1966. P. 213-214. Позже, в том числе под влиянием Деннета, Куайн несколько смягчил свою позицию — Quine W. V. *Pursuit of Truth*. Cambridge MA, 1990. P. 71-72. См. также: Quine W, V. *States of mind // The Nature of Mind*. P. 287-288. Здесь Куайн дает понять, что симпатизирует «аномальному монизму» Д. Дэвидсона, т. е. сближается с одним из вариантов теории тождества.

⁶ Dennett D. C. *An overview of my work in philosophy*. TS 1996.

⁷ См.: Dennett D. C. *Brainchildren: Essays on Designing Minds*. L., 1998. P. 357.

⁸ Неудивительно, что в студенческие годы Деннет критиковал куйановский тезис неопределенности перевода — *ibid*.

⁹ Dennett D. C. *Re-introducing The Concept of Mind // Ryle G. The Concept of Mind*. Chicago, 2002. P. XII.

¹⁰ См., напр.: Ryle G. *The Concept of* P. 62, 245, 280-281.

¹¹ «Таких объектов, как ментальные картины, вовсе не существует» — *ibid.* P. 254.

¹² См.: Sellars W. *Empiricism and the Philosophy of Mind*. Cambridge MA, 1997.

¹³ Ср.: *ibid.* P. 13.

¹⁴ См.: Blackmore S. *Consciousness*. P. 411-414.

¹⁵ Р. Соренсен на простом примере показывает, что даже сильным интуициям не всегда можно доверять. Представим ленту, опоясывающую земной шар по экватору. Теперь удлиним ее на 30 см. Насколько она будет отступать от земли? Интуиция уверяет: на крошечную величину. В действительности — почти на 5 см. См.: So-rensen R.A. *Thought Experiments*. N. Y., 1992. P. 31.

¹⁶ См., напр., «эксклюзивную» статью Рорти: Рорти Р. Бум в исследованиях сознания в свете элиминативизма // *Историко-философский альманах: Вып. 2. М., 2007. С. 242-244.*

¹⁷ Самые известные элиминативисты — супруги Пол и Патрисия Черчленды — пытались воспитывать своих детей в духе элиминативизма, избегая употреблять о них ментальные термины. Предлагаю читателям догадаться с одного раза, успешным ли оказалось их смелое начинание.

¹⁸ См.: Ryle G. *The Concept of Mind*. P. 62.

¹⁹ Этот термин используется в качестве синонима «философского бихевиоризма».

²⁰ Ryle G. *The Concept of Mind*. P. 329-330.

²¹ *Ibid.* P. 323.

²² *Ibid.* P. 200, 240.

²³ Хотя, как мы уже знаем, с разной степенью уверенности. ²⁴ См.: Dennett D. *Content and Consciousness*. L., 1986 (1969). P. 4-6.

²⁵ См. гл.2.

²⁶ Interview with Daniel Dennett conducted by Bill Uzgalis in Boston, Massachusetts on December 29, 2004 // *Mind and Machines* 16:1 (2006). P. 7-19.

²⁷ Dennett D. *Content and Consciousness*. P. 17.

²⁸ *Ibid.* P.8-15.

²⁹ Dennett D. *Consciousness Explained*. Boston, 1991. P. 457. В письме автору от 4 августа 2004 г. Деннет уточнял, что он «использует теорию интенциональности для того, чтобы показать, как человеческое сознание возникает в качестве специфического феномена внутри этой теории».

³⁰ Dennett D.C. *Brainchildren*. P. 363.

³¹ *Ibid.* P. 364.

³² *Ibid.*

³³ *Ibid.*

³⁴ См.: Dennett D. *Content and Consciousness*. P. 122.

³⁵ Dennett D. *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge MA, 1981 (1978). P. 4-6. ³⁷ Dennett D. *Intentional systems theory*. TS 2007. P. 2.

³⁷ Замечательный рассказ Дианы Гаспарян о туссанской конференции 2006 г., а также о том, что ей предшествовало, опубликован на сайте <http://www.histphil.ru>. От себя могу добавить три вещи.

Во-первых, нельзя не отметить подготовительную работу организаторов. Они тщательно проинструктировали докладчиков, и в результате пленарные заседания напоминали показательные бои или отредактированные телевизионные шоу — в хорошем смысле. Во-вторых, лучшее время получали не именитые авторы, или не только они, но и просто те, кому было что сказать. Хотя новички не могли тягаться с Дугласом Хофштадтером, утонченно рассуждавшим о «странных петлях» обратной связи в мозге, образующих человеческое Я. В-третьих, конференция незаметно перетекала в разного рода художественные перформансы. Кульминацией же стала традиционная финальная вечеринка «Конец сознания», а кульминацией вечеринки — следующая сцена, которую, к сожалению, смог по-настоящему прочувствовать только я. Вообразите веселую толпу, выстроившуюся в гнутую очередь за фуршетной выпивкой — гомон, смех... И вдруг рядом со мной возникает захмелевший человек, обводит очередь мутным взглядом и выдыхает «странная петля...». Но настоящий конец сознания наступил, когда я слегка спиил Сью Блэкмор и, разглядывая ее многоцветные волосы, болтал с ней о Деннете и буддизме...

³⁸ Очевидно, таким образом, что интенциональная установка является продолжением или даже «подвидом» дизайнерской — см.: Dennett D. C. *Intentional systems theory*. P. 3.

³⁹ Dennett D. C. *Brainstorms*. P. 12.

⁴⁰ См.: Dennett D. C. *Darwin's Dangerous Idea: Evolution and the Meanings of life*. L., 1995. P. 370-383.

⁴¹ Dawkins R. *The Selfish Gene*. Oxford, 1999 (1976). P. 192. Подробный анализ различных аспектов теории мемов — см.: Blachmore S. *The Meme Machine*. N. Y., 1999.

⁴² См.: http://fragmeiits.consc.uet/djc/2005/01/books_in_the_ph.html.

⁴³ Dennett D. C. *Consciousness Explained*. P. XI.

⁴⁴ *Ibid.* P. 16.

⁴⁵ Деннет, впрочем, не исключает, что гетерофеноменологический метод может быть распространен и на изучение психики животных. Обсуждение этого вопроса — см.: Elton M. *Daniel Dennett: Reconciling Science and Our Self-Conception*. Cambridge, 2003. P. 171 -173.

⁴⁶ Аутофеноменология, пишет Деннет, «ведет к искушению: искушению рассматривать наши убеждения от первого лица не в качестве данностей, а в качестве непреложной истины» — Dennett D. C. *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. Cambridge MA, 2005. P. 113.

⁴⁷ К примеру, Деннет пытается доказать, что цвета (эталонные квалиа) являются не какими-то субъективными данностями, а объективными свойствами отражающих поверхностей.

⁴⁸ Dennett D. C. *Consciousness Explained*. P. 458: «В конце концов, главный источник мифа о Картезианском театре — леностная экстраполяция интенциональной установки на целое. Рассмотрение сложного, подвижного единства как единого агента — впечатляющий способ усматривать паттерн в совокупной деятельности; эта тактика естественна для нас и, возможно, даже поощряется генетически в качестве способа восприятия и мышления».

⁴⁹ Ibid. P. 33-37. ⁵⁰ Ibid. P. 107.

⁵¹ Неудивительно, что Деннет не жалуется Крика и Коха и редко ссылается на них. А в «Объясненном сознании» он сам говорит, что они близки «картезианскому материализму» — *ibid.* 256.

⁵² Dennett D. C. *Consciousness Explained*. P. 215.
⁵³ *Ibid.* P. 211.

⁵⁴ *Ibid.* P.210.

⁵⁵ Ср.: Kirk R. "The best set of tools"? Dennett's metaphors and the mind-body problem // *The Philosophical Quarterly* 43: 172 (1993). P. 338.

⁵⁶ Zawidzki T. Dennett. P. 77. Отметим, что Завидский не проводит различий даже между основными вариантами деннетовской теории сознания.

⁵⁷ Свою роль могла сыграть и философская критика. К примеру, П. Черчленд доказывал, что «виртуальная машина», состоящая из некоего набора правил, не может работать в мозге с его коннекционистской архитектурой. См.: Churchland P. M. *Densmore and Dennett on virtual machines and consciousness* // *Philosophy and Phenomenological Research* 59: 3 (1999). P. 763-767. Сила этого возражения зависит от того, действительно ли коннекционистские паттерны в мозге нельзя интерпретировать на языке символьных конструкций, как склонен считать Черчленд.

⁵⁸ Ср.: Koch C. *The Quest for Consciousness*. P. 269-277; Leopold D. A. *Logothetis N. K. Activity changes in early visual cortex reflects monkeys' percepts during binocular rivalry* // *Nature* 379 (1996). P. 549-553.

⁵⁹ А именно данных, показывающих аналогию между поведением животных и «сознательным» поведением человека. Что же касается феномена бинокулярного соперничества, то в работах Деннета трудно отыскать хотя бы упоминание о нем, что само по себе достаточно любопытно. Вариант того, что он не знал о нем, должен быть исключен. Ведь об этом шла речь хотя бы в статье Ф. Крика и К. Коха 1990 г.; которую он цитирует уже в «Объясненном сознании».

⁶⁰ Dennett D. C. *Sweet Dreams*. P. 172. Ср.: Dennett D. C. *Two steps closer to consciousness* // Paul Churchland, ed. by V. Keeley. N. Y., 2006. P.193-209.

⁶¹ В «Сладких грезах» Деннет утверждает, что концепция набросков была недостаточно иллюстративной. Но в нашей беседе (см. ниже) он заметил, что отказался от этой метафоры потому, что набросок как таковой предполагает существование того, кто читает его, и в этом смысле она наводит на мысль, что в мозге все же существует какой-то Наблюдатель. Любопытно, однако, что в «Сладких грезах» Деннет высказывал подобные соображения и о метафоре «славы в мозге», предлагая заменить ее на метафору «влияния» или «фантазийного эха» — Dennett D. C. *Sweet Dreams*. P. 136, 160. Но к 2007 г. он, похоже, вновь вернулся к метафоре «славы в мозге».

⁶² *Ibid.* P. 131-132.

⁶³ Впрочем, Баарс не отрицает, что он, скорее, обобщил еще более ранние теории.

⁶⁴ В статье «Реальное сознание» (1994), перепечатанной в сборнике «Дети мозга», было зафиксировано некое переходное состояние — см.: Dennett D. C. *Brainchildren*. P. 131-139.

⁶⁵ Показательно, что Баарс, рассуждая о своей теории, стал использовать деннетовскую метафору «славы в мозге». См.: Baars V. *The Global Workspace Theory of Consciousness*. P, 240.

⁶⁶ Dennett D. C. *Consciousness Explained*. P. 257.

⁶⁷ *Ibid.* P. 254.

⁶⁸ *Ibid.* P. 253.

⁶⁹ *Ibid.*

⁷⁰ Кроме того, если речь идет о коротких промежутках времени, то в некоторых случаях — когда, к примеру, происходит быстрая последовательная экспозиция двух предметов и человек утверждает, что он вообще не видел первого предмета, — мы, по Деннету, не в состоянии выбрать между двумя вариантами объяснения этого феномена, «сталинистским» (первый предмет вообще был не допущен до сознания и все ограничилось восприятием второго) и «оруэллианским» (предмет был осознан, но память об этом событии была стерта). Сама попытка выбрать между этими сценариями предполагает, считает Деннет, ложное понимание потока сознания как чего-то реального. Критики Деннета, и в частности, Н. Блок, однако, указывали, что можно представить вполне конкретные и реальные нейробиологические эксперименты, которые позволят сделать выбор. См.: Block N. Review of Daniel Dennett, *Consciousness Explained* // Block N. *Consciousness, Function and Representation: Collected Papers, Volume 1*. Cambridge MA, 2007. P. 136.

⁷¹ Ср.: Dennett D. C. *Consciousness Explained*. P. 275.

⁷² Ср.: Dennett D, C. *Sweet Dreams*. P. 132. Разумеется, никакая трансляция не состоится, если она не продлится какое-то минимальное время. Но Деннет имеет в виду гораздо большие отрезки.

⁷³ *Ibid.* P. 162-163. См. также: Dennett D. C. *Brainchildren*. P. 138.

⁷⁴ Это обстоятельство недостаточно учитывает С. Шнейдер — см.: Schneider S. Daniel Dennett on the nature of consciousness // *The Blackwell Companion to Consciousness*. P. 318-319.

⁷⁵ Ср.: Dennett D. C, Akins K. Multiple drafts model // *Scholarpedia* 3:4 (2008). P. 4321.

⁷⁶ См.; Hameroff S. The good, the bad and the octopus // *Journal of Consciousness Studies* 14: 8 (2007). P. 106.

⁷⁷ Blackmore S. Destroying the zombic hunch. Review of *Sweet Dreams* by Daniel Dennett // *Nature* 5 (2005) P. 21-23.

⁷⁸ См.: Blackmore S. Driving beyond consciousness: a test of multiple drafts theory// *Toward a Science of Consciousness* — Tucson 2006, P. 112-113. Д. Б. Волков (рукопись) предположил, что результаты эксперимента Блэкмор можно истолковать так, что теория Деннета не окажется фальсифицированной — при допущении того, что субъекты, принимавшие участие в эксперименте, периодически совершали аутопробы, нивелировавшие результаты внешних «проб». С этим можно согласиться, но ценой такого допущения окажется то, что предсказательный потенциал теории Деннета, который и так неочевиден, может снизиться чуть ли не до минимума. Это особенно

болезненно для него, так как он считает свои построения эмпирическими гипотезами.

⁷⁹ Dennett D. C. *Consciousness Explained*. P. 240.

⁸⁰ Ibid. P. 228.

⁸¹ Ibid. P. 406.

⁸² Ibid. P. 310.

⁸³ Все началось с того, что на каком-то из семинаров Кирка это слово употребила одна наивная (или не очень наивная) первокурсница — см.: Kirk R. *Zombies and Consciousness*. N.Y., 2005. P. 25. Поначалу Кирк доказывал, что зомби логически возможны — см.: Kirk R. *Sentience and Behaviour // Mind* 83 (1974). P. 43-60. Позже он стал решительным критиком этой идеи.

⁸⁴ Dennett D. C. *Consciousness Explained*. P. 73.

⁸⁵ Bennett D. C. *Sweet Dreams*. P. 16.

⁸⁶ Ibid. P. 23.

⁸⁷ Ibid.

⁸⁸ При этом, как заметил Деннет в одном интервью, с позиции третьего лица мы могли бы лучше знать, «каково быть» этими людьми, чем было бы «интроспективно» известно им самим.

⁸⁹ Сам Деннет, однако, считает, что мы приписываем интенциональные состояния, а значит, и сознание в том числе и сами себе. Это следует из его концепции множественных набросков начала 90-х гг., и это взгляд, который по-прежнему близок ему. По крайней мере, он заявил об этом в нашей кембриджской беседе, которая, кстати, получилась довольно любопытной, как и встреча с Н. Блоком — одним из самых креативных умов в современной аналитической философии, — состоявшаяся в тот же день. Поэтому я решил опубликовать заметки о них, сделанные по горячим следам. На этих встречах я был вместе с Д. Б. Волковым. — 3 апреля 2007 г., Кембридж, Массачусетс. Добралась с Д. В. из Нью-Йорка на машине, под руководством мудрой спутниковой навигаторши. Несколько раз мы ее, правда, неверно понимали. Дорога не слишком запоминающаяся. Только в Коннектикуте неплохие виды — красивые каменные осыпи на всхолмиях. И еще романтичный скайлайн Бостона на подъезде к Кембриджу. Сам Кембридж выглядит совершенно не по-американски. Красные каменные европейские домики (хотя кое-где попадаются и стеклянные страшилища образца где-то 60-х), узкие улочки. Великолепный книжный магазин. Море философской литературы. Приехали около полудня. Оставили машину на стоянке и пешком добрались до Sandrine's Bistro, где на час дня и была назначена встреча с Блоком (всего она продолжалась чуть менее двух часов). Дизайн этого ресторана явно недоработан, из общей картины выбивались какие-то трубы под потолком и лампы в виде фар. У нас был еще час, и мы решили перекусить. В 13.00 у входа появился Блок, и мы его сразу взяли в оборот, усадили—и приступили. Нас интересовали философские вещи, но Блок все время делал попытки, хотя и не очень настойчивые, свернуть в сторону. Спрашивал, почему нам так нравится Нью-Йорк, чем мы занимаемся (в ответ на этот вопрос я подарил ему свою английскую статью о сознании). Охотно рассуждал о достоинствах Нью-Йоркского университета, о политике приглашения звезд, которую начал еще Rutgers University, о том, что Нью-Йоркский университет предоставляет де-

шевые квартиры, и т. п. Восхищался своими аспирантами, о студентах говорил осторожнее. Много слов было сказано о его статье в BBS, на которую в ближайшее время будет написано множество комментариев — этот журнал регулярно делает подобные вещи. Выглядит Блок экстравагантно. Похож на бабуся. Высокий, длинные седоватые волосы и светящиеся голубые глаза, то и дело обретающие презабавнейшее выражение. Смешной. Но философские идеи подавал суховато. Заметил, кстати, что обычно он не соглашается на интервью (хотя у нас, конечно, это было не интервью) — и поэтому, в частности, уже неоднократно отказывался от выступления в Philosophy Talk (сам слушал эту передачу всего пару раз, хотя и считает ее хорошей). Много рассуждал о «провале в объяснении» и о невозможности нынешними средствами решить «трудную проблему сознания». Но настаивал, что в будущем прогрессирующая нейронаука может сделать возможным ее решение. Как его можно мыслить, мы пока не знаем. Но должен быть союз аналитической философии и науки. Аналитическая философия переживает расцвет, хотя сама она изменилась, перестала быть чисто концептуальной. Не слишком ясна его позиция относительно того, где проходит грань между концептуальными и эмпирическими вопросами и методами. Отказываясь братья за «трудную проблему», тем не менее заметил, что является физикалистом. Квалиа имеют физические свойства. В этом физикалистском ключе, сближающем его с теоретиками тождества (но, как он подчеркнул, не в функционалистском варианте Смарта, Льюиса и Армстронга) решает и проблему ментальной каузальности. Я спросил его о том, не считает ли он, что тезис о тождестве качественного и физического в принципе не допускает верификации и поэтому ложен, на что он ответил — как в свое время и Маккинси, — что верификация здесь не обязательна и что здесь работает «принцип лучшего объяснения». Корреляция в любом случае недостаточна. Как недостаточно говорить о корреляции, к примеру, теплоты и определенного движения молекул (я не помню точно, именно ли этот пример он приводил), а более продуктивно говорить, что теплота и есть это движение, так и в случае объяснения сознания лучше отождествлять квалиа с физическими процессами — хотя детали нам неизвестны из-за все того же провала в объяснении. На вопрос о функциях квалиа ответил сначала, что они не вполне ясны, хотя потом добавил, что они как-то влияют на формирование access consciousness, если я не путаю. Много говорили о Китайской нации и Китайской комнате. Он уверен, что Сёрл действительно заимствовал идею у него (удивился упоминанию рассказа Сёрла о самолете, я также заинтересовался, узнав, что и Деннет предлагал нечто подобное). Хотя тут же он заметил, что аргумент Сёрла о понимании, а его — о сознании. Впрочем, он согласен, что и аргумент Сёрла в конечном счете о сознании. Считает, что аргумент Китайской комнаты сохраняет свою силу, хотя он всего лишь апеллирует к интуициям. Обращаться к ним можно и даже нужно — если мы, к примеру, хотим выяснить, кто в конечном счете прав, он, скажем, или Деннет, хотя и нельзя безусловно доверять этим интуициям. Я спросил его об изменении его взглядов, о том, что он стал более осторожен в своих оценках соотношения феноменального сознания и сознания доступа. Он сказал, что его позиция по сути не изменилась, но чуть позже, говоря о функциях квалиа, добавил, что его взгляды на

этот счет еще не устоялись. Д. спросил его о том, что, собственно, понимается под «сознанием доступа». Кто, собственно, должен осуществлять доступ? Он ответил, что Я, но само это Я есть не более чем система. Я спросил его об истоках когнитивной науки, поскольку они как-то не очень прозрачны. Он предложил различать когнитивную революцию, связанную с именами Хомского и Патнэма, и институционализацию когнитивной науки, произошедшую на рубеже семидесятых—восемидесятых годов благодаря деятельности «Слоан Фаундейшн». Рассказывал, что они решили не делать специальную степень по когнитивным наукам, так же как ее не делают сейчас по Consciousness Studies — исходя из соображений трудностей отыскания работы для выпускников, насколько я понял (хотя это и парадоксально). Но он считает, что на каждой кафедре психологии должны быть люди, ведущие тему сознания. Я спросил его об учителях. Он учился в Гарварде и очень тепло вспоминал о Патнэме (но сейчас он с ним почти не общается). Патнэм казался человеком, который знал все и мог компетентно говорить на любую тему. Исключительно притягательная личность. А вот Куайн не нравился ему. Он хотел обсуждать только собственные идеи. Упомянул о влиянии Хомского, хотя он и не учился у него. Порекомендовал нам сходить на публичную лекцию французского нейрочеловека Деана (уточняя его фамилию — на слух было неясно, о ком он говорит, - написал ее прямо на скатерти), заметив между тем, что идеи Деана противоположны его собственным — хотя он не раз ссылается на его результаты в статьях. В этой связи очень тепло отозвался и о Кристофе Кохе. Узнав уже на улице (предварительно расплатившись за всё, несмотря на наши вялые протесты), что в пять часов мы будем встречаться с Деннетом в ресторане ОМ, сказал, что вы «услышите совсем другие истории»... Мы с Д. отправились на лекцию Деана в огромном такси с угрюмым темнокожим водителем. Лекция была в одном из старых корпусов Гарвардского университета, расположенном рядом с самым высоким зданием в Кембридже (этажей 17). Пришли минут за десять. Заканчивалась предыдущая лекция, и все толпились в фойе. Удобная аудитория, чем-то напоминающая наши. Пришел Блок. Вдруг появился и Деннет, который, не замечая нас, сел рядом со мной — то и дело поглядывая на часы, чтобы не опоздать на встречу с нами. Лекция была посвящена осмыслению данных fMRI по участкам мозга, отвечающим за чтение. Загадка в том, что, хотя генетически эти участки не могут быть детерминированы, у всех людей они одни и те же. Как такое может быть? Гипотеза Деана состоит в том, что в этих участках произошло перераспределение функций. Изначально они отвечали за распознавание элементарных геометрических форм. И из этих форм нетрудно составить чуть ли не все алфавиты. В пять часов Деннет вышел, и мы за ним. Поздоровались и направились в ОМ - гастрономический ресторан, где нам предстояло отведать немало изысканных блюд, ни названий которых, ни их ингредиентов я, разумеется, не запомнил. По пути Деннет рассуждал о формировании в мозге различных систем распознавания, но потом — поскольку из-за моей манеры ходить мы шли очень быстро — попросил чуть сбавить ход. Сказал, что чувствует себя после перенесенной операции в общем неплохо, занимается с инструктором в тренажерном зале (был там с утра и в этот день), поднимает гири и т. и. Рассказывал о том, как

проходила операция, как охлаждали его мозг — до этого он уже сообщал подробности прессе, так что мы многое знали и сами. Боялся, не приведет ли все это к каким-то нарушениям, и поэтому сразу после операции попросил жену принести ему кроссворды — и, только решив их с прежней скоростью, успокоился. В ресторане сначала долго определяли, как и что мы будем есть. Затем приступили к раздаче подарков и фотосессии. Спросил, я ли написал главу о нем в нашем учебнике, и узнав, что да, попросил подписать ее, а потом и сфотографировался с учебником в руках... Деннет произвел очень сильное впечатление. Это мощный и оригинальный ум. В нем есть что-то от иллюзиониста или гипнотизера — сам Деннет любит цитировать в этой связи Б. Вурхеса, назвавшего его дьяволом, потому что он, так сказать, крадет у людей душу, оставляя телесные одежды без облаченного в них короля, т. е. считая сознание иллюзией. Большую часть беседы, которая продолжалась около трех часов, с пяти до восьми вечера, он монотонно говорил, словно суггестируя свои идеи, и периодически сопровождал свои слова размеренными пассажами. Беседу мы начали с его вопроса о наших занятиях, в ответ на который было сказано, что я пишу книгу «Трудная проблема сознания». Деннет заявил, что считает, будто такой проблемы не существует, и я подтвердил, что, конечно, знаю о его позиции. Меня интересовало несколько вопросов, и я получил на них ответ. Во-первых, я хотел узнать о том, как он воспринимает эволюцию собственных идей. Он сказал, что глобальных изменений не было. Теорию «славы в мозге» он не считает в чем-либо по сути противоречащей теории «множественных набросков». Недостаток более раннего названия, сказал он (я не припомню, чтобы он где-то писал об этом, — хотя, как я впоследствии установил, подобное замечание делал по поводу его теории Р. Кирк еще в 1993 г.), в том, что слово «набросок» неявно предполагает существование того, кто его читает. И это нежелательный образ. «Слава» более нейтральная метафора. Я спросил его (памятуя о своем диспуте с Гиренком), зачем вообще использовать метафоры. Ответ свелся к тому, что метафоры — это просто высокоуровневые описания низкоуровневых процессов. Перешли в этой связи к проблеме метафоры сознания как software. Я спросил, не кажется ли ему, что software как таковое есть просто абстракция. Реально оно вносит изменения в физические каузальные цепи — и испаряется. Он не вполне согласился с этим (хотя в своей поздней трактовке «виртуальных машин» в мозге он занимает примерно такую же позицию)... Рассуждал о феноменологии, о том, что если разные люди будут описывать, кажется, взаимоотношения внутри какой-нибудь бейсбольной команды, то эти описания будут различаться — но это, и только это, и есть феноменология, пусть и бихевиористичная (но он не видит в этом ничего плохого). Говорил об интенциональной установке, о том, что нет различия между производной и изначальной интенциональностью, т. е. о том, что человеческая якобы изначальная интенциональность тоже производна, и производит ее естественный отбор. Я сказал, что производность здесь берется в разном значении, но Деннет с этим не согласился. Тогда я спросил, считает ли он, что мы приписываем интенциональные состояния самим себе, и он ответил, что да, что эти интенциональные состояния есть полезная иллюзия. Тут Д. задал один из ключевых вопросов, который я и сам в свое время задавал Деннету в письме — и не

получил ответа: как понимать иллюзию? Каков ее статус? Отвечая, Деннет сказал нечто новое для меня. Он сослался на то, что иллюзии могут быть и у роботов и что в свое понимание иллюзии он не вкладывает ничего большего. В связи с проблемой реальности сознания не могли не поговорить об интерпретациях квантовой механики. Я знал, что Деннет отвергает интерпретации, связанные с признанием сознания как фактора в коллапсе волновой функции. Деннет заметил, что ему не близка и копенгагенская интерпретация, так как она придает большое значение измерению (и косвенно ведет к признанию ключевой роли сознания в этом коллапсе). Потом он упомянул недавнюю книгу одного нобелевского лауреата, в которой дана устраивающая его пробабилистская интерпретация — а устраивает она его, видимо, потому, что перекликается с его теорией личности и критикой субстанциального дуализма. Тут, кстати, я напомнил ему, что в одном месте он как-то сказал, что не считает дуализм логически противоречивой теорией. Но если так, то почему он все-таки сходу отвергает дуализм. Отвечая, он повторил то, что говорит всегда, а именно то, что дуализм — пораженческая теория, так как он не предлагает никакого объяснения взаимодействия сознания и мозга. Тут же он упомянул темную материю, заметив, что вот дуализм обычной материи и темной материи — нормальный дуализм, в отличие от дуализма материи и сознания. В ответ на это я сказал, что мы точно так же не понимаем принципов взаимодействия темной и обычной материи, как не понимаем принципов взаимодействия мозга и сознания. Это замечание было несколько неожиданным для Деннета. Он сказал, что это интересно, и что он никогда не думал об этом, и что, возможно, у него имеются предубеждения на этот счет. Но он полагает, что различия есть (любопытно, что Деннет, похоже, не сомневается в реальности темной материи, хотя некоторые считают, что это не настоящая, а мнимая сущность, продукт некорректного *ad hoc* допущения). Вернувшись к особенностям его новой теории сознания и изменений в его взглядах, я спросил его о том, не является ли явным изменением то, что он стал гораздо более осторожно говорить о культурной обусловленности сознания, чем раньше. Этот вопрос вывел на тему сознания у животных. Деннет признал, что раньше он, возможно, был излишне радикален, отрицая его. Новейшие эксперименты могут заставить его изменить свое мнение. Но к оценке этих экспериментов надо подходить очень осторожно, заметил он, и привел несколько иллюстраций. Сначала он прокомментировал знаменитые эксперименты с шимпанзе, якобы распознающими себя в зеркале. Убедительных доказательств долгое время не было, а те видеозаписи этих экспериментов, которые он смотрел, были совершенно невыразительными. И лишь совсем недавно появились такие записи, которые не оставляют на этот счет сомнений (он изобразил шимпанзе, продемонстрировав, как все это выглядит), — надо, правда, помнить, добавил он, что лишь три из десяти шимпанзе способны узнавать свое отражение. Еще более интересный случай был связан с экспериментами, вроде бы демонстрирующими развитые коммуникативные способности обезьян. Деннет слушал доклад об этом на одной из конференций. Доклад был настолько убедительным, что смотреть видеозапись экспериментов не было, казалось, никакой необходимости. И тем не менее он решил сделать это, пожертвовав ланчем. Результат был

ошеломляющим: он совершенно не мог увидеть этой коммуникации, глядя на экран и слушая комментарии исследовательницы. Она вкладывала в увиденное то, чего там совершенно точно не было. После этого Деннет твердо решил не доверять словам, когда речь идет об интеллектуальных способностях животных. Кажется, в этой связи он сказал, что одна из задач философии — поддерживать такой критический настрой. И тут мы перешли к обсуждению того, чем вообще должна заниматься философия. Деннет настаивал, что философия — это деятельность, помогающая ставить правильные вопросы в тех областях, где еще неясно, как ставить эти правильные вопросы. Но ответы на них в любом случае дает не она, на них отвечает наука. Говорил, что эта особенность философии может вгонять в депрессию тех людей, которые ценят точность и наукообразность... За окном стемнело, и беседа вдруг переменялась. Произошло что-то неувидимое, Деннет словно преобразился в одного из «Таинственных людей» — и предложил нам сыграть в удивительную игру, зная алгоритм которой можно почти наверняка выигрывать. Он выиграл у меня пару раз, а потом объяснил алгоритм — но я уже почти ничего не понимал, одурманенный винами, и даже признался, что, бездумно повторяя тот алгоритм, чувствую себя человеком в Китайской комнате. А потом был Пазл и гномы, спасающие друг друга... Затем Деннет и вовсе на какое-то время растворился в воздухе. Расплатившись и вновь обретя его, мы вскоре тепло распрощались на промозглой кембриджской улице. Деннет (пообещавший нам перед тем, как расстаться, еще в ОМе, что его новая книга тоже будет о сознании; и мы долго размышляли над названием: он шутливо предложил «Действительно объясненное сознание» — мой вариант был *Consciousness Better Explained* — и добавил, что люди запоминают заглавия его книг) отправился домой, чуть ли не пешком, а нам предстояла долгая ночная дорога в Нью-Йорк. Мы попали в пробку на 1-95, и к тому же наша электронная подружка неправильно поняла нас и отправила куда-то в пригороды нашего мегаполиса: на родной *Washington Square*, мы очутились лишь в 2 часа ночи.

⁹⁰ Jackson F. Epiphenomenal qualia // *Philosophical Quarterly* 32 (1982). P. 127-136.

⁹¹ Jackson F. Mind and illusion // *There's Something About Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*, ed. by P. Ludow, Y. Nagasawa, and D. Stoljar. Cambridge MA, 2004. P. 421-442.

⁹² К примеру, способность «узнавать, воображать и вспоминать данное состояние» — *ibid.* P. 439.

⁹³ См.: Papineau D. *Thinking about Consciousness*. N. Y., 2004 (2002). P. 47-72. Отметим, что все эти варианты проходят лишь при предположении, что субъективные данности, подобные данности красного, могут быть тождественны физическим процессам в мозге. Иными словами, мысленный эксперимент Джексона не опровергает физикализм лишь при условии осмысленности теории тождества. Мы видели, однако, что у этого предположения нет резонансов. Теория тождества опирается на неverified допущение. Но если согласиться с тем, что она ошибочна, опровергнуть физикализм можно проще, чем это делал Джексон. (1) Вообразим, к примеру, лимон. В нашем сознании имеется визуальная данность, нечто желтое. (2) Так как речь

идет о воображении, бессмысленно искать эту данность среди внешних объектов. (3) Очевидно, однако, что и визуально изучая свой мозг с позиции третьего лица, мы не найдем в нем ничего желтого, и уж точно не найдем там лимона. (4) Следовательно, данности третьего лица, т. е. физические данности, не исчерпывают всех данностей, что, собственно, и означает, что физикализм ложен.

⁹⁴ См.: Dennett D. C. *Consciousness Explained*. P. 398-401; Dennett D. C. *Sweet Dreams*. P. 103-129.

⁹⁵ Ср.: Churchland P. M. *Knowing qualia: a reply to Jackson (with Postscript 1997)* // *There's Something About Mary*. P. 164-167.

⁹⁶ На Туссанской конференции 1996 г. Д. Чалмерс направил на Деннета сконструированный им «прибор», похожий на фен и предназначенный для индикации сознания. Лампочка не загорелась.

⁹⁷ См.: Dennett D. C. *Sweet Dreams*. P. 10-13. Здесь Деннет говорит о том, что даже если бы оказалось, что благодаря особым свойствам нейронов каждый из них обладает неким «квантом» сознания, из сложения которых и возникает наше субъективное Я, то это еще не дало бы ответа на главный, по мнению Деннета, вопрос о том, по каким принципам они должны быть сложены. Из этих пассажей следует, что Деннет попросту выносит за скобки вопрос о субъективности.

⁹⁸ *Ibid.* P. 12.

⁹⁹ Так что Деннет напрасно стыдился того, что он является представителем той же профессии (т. е. философии), что и К. Макгинн, отрицающий возможность объяснить сознание. Вообще, Деннет на удивление много рассуждает о том, как ему неуютно среди коллег по профессии.

¹⁰⁰ Одна из формулировок «трудной проблемы»: почему мы не зомби?

¹⁰¹ Деннет упоминает фокус «Умная колода» (*The Tuned Deck*) Р. Халла — см.: Dennett D. C. *Sweet Dreams*. P. 72-75. Думаю, не менее удачный пример — «Лазерная иллюзия» Д. Копперфилда.

¹⁰² См.: <http://ru.youtube.com/watch?v=wIdxbJyvfNw&feature=related> и др. части этой лекции. Ср.: Dennett D. C. *Consciousness Explained*. P. 455.

Глава 4 ЧАЛМЕРС: ВСЕ РЕШЕНИЯ ПЛОХИ...

В предыдущих главах мы уже не раз сталкивались с идеями Чалмерса. И в этом нет ничего удивительного: в начале XXI века, пожалуй, не удастся найти ни одной значимой книги о сознании, в которой не упоминались бы работы этого автора¹. Феномен этот, конечно, требует толкования, и самое простое объяснение состоит в том, что Чалмерс — очень глубокий и оригинальный мыслитель. Но нельзя отрицать и некоторые другие факторы, к примеру, то, что он оказался в нужное время в нужном месте. Его туссанское выступление 1994 г., в котором он различил «легкие» и «трудную» проблемы сознания, обеспечило ему мгновенную славу, закрепленную фундаментальным трудом 1996 г. — «Сознающий ум», в котором он попытался найти средний путь между установками Деннета и Сёрла, объединив их позитивные интуиции. В год публикации этой книги Чалмерсу исполнилось всего 30 лет. Нетрудно догадаться, что в то время он лишь начинал свою философскую карьеру, тем более что его первые шаги в академической среде вообще не были связаны с философией. Он стартовал как математик в своей родной Австралии (в университете Аделаиды)², затем продолжил математическое образование в Оксфорде, но чем больше он занимался математикой, тем сильнее хотел поменять специальность. В этом ему помог знаменитый Дуглас Хофштадтер, ставший его научным руководителем в Университете Индианы (США), где Чалмерс и получил докторскую степень по философии в 1993 г.³ — накануне той самой Туссанской конференции, которая принесла ему мировую известность. Поработав в 1993—1998 гг. в Вашингтонском (Сент-Луис) и Калифорнийском (Санта-Крус) университетах, в 1999 г. Чалмерс обосновался на кафедре философии Аризонского университета в Туссане и возглавил там Центр исследований сознания. В 2004 г. он вернулся на родину, заняв должность профессора Австралийского национального университета (Канберра) и директора Центра сознания.

Из биографической справки Чалмерса видно, что он не чужд организаторской деятельности. Возможно, его влияние в наши дни отчасти связано и с его успехами на этом поприще. Он, кстати, умеет превращать академические мероприятия в настоящие шоу, в которых иногда и сам солирует, исполняя свой сумасшедший «Зомби-блюз»⁴. Чалмерс, кроме того,

курирует оксфордскую серию по философии сознания, и за последние годы под этой рубрикой вышло немало качественных книг. Все эти вещи, как шутит Дэвид, позволяют ему находить оправдания для того, чтобы не заниматься «реальной работой». Между тем именно эта работа — которую он время от времени все же делает — представляет главный интерес для нас.

Проблема сознания увлекла Чалмерса еще в студенческие годы, когда он изучал математику и computer science в своем первом университете. Позже, переключившись на философию и сосредоточившись на теме сознания, он, однако, не забыл прошлое: уже в ранних рукописных текстах Чалмерса (опубликованных им на своем веб-сайте) мы видим стремление сочетать «деннетовский» компьютерно-функционалист-ский подход к изучению работы мозга с «сёрловским» признанием реальности качественных состояний.

Особый интерес в этой связи представляет набросок статьи «Точки зрения от первого ПИ ОТ третьего лица», созданный в 1989 г., когда Чалмерс только-только поступил в аспирантуру Индианы. Буквально с первых строк он заявляет о «великой тайне» субъективности: «знаменитая проблема сознание—тело в наши просвященные материалистические времена сводится исключительно к вопросу: что такое "первое лицо" и как оно возможно?»⁵ Обратим внимание, что Чалмерс пока еще пользуется старым термином «проблема сознание—тело». Вместе с тем в тексте неоднократно звучит мысль, что эта формулировка устарела и больше не является адекватной — так готовится почва для появления на свет «трудной проблемы сознания».

Старая формулировка, полагает Чалмерс, устарела по крайней мере по двум причинам. Во-первых, термин «сознание» (mind), некогда обозначавший субъективную реальность «от первого лица», в наши дни оказался приватизирован когнитивной наукой, проводящей линию на изучение ментальных феноменов с объективистских позиций компьютерной метафоры, с точки зрения «третьего лица». Во-вторых, в современной редакции проблемы сознание—тело речь идет, скорее, не о теле или телесности вообще, а о мозге. Так или иначе, но за неимением лучшего Чалмерс пока пользуется старой терминологической конструкцией. Не применяет он и термина «зомби». Но сама идея «существа, которое, как кажется, обладает всеми атрибутами от третьего лица, необходимыми, скажем, для сознания, но которое вместе с тем лишено "того, кто переживает" от первого лица», обсуждается им. Выражение «как кажется», акцентированное Чалмерсом, выделено им не случайно: он склоняется к выводу, что в действитель-

ности зомби немислимы и невозможны (это убеждение было у него еще в оксфордский период). Иначе говоря, он считает, что "первое и третье лицо неразделимы", и обещает показать, почему дело обстоит именно так. Там же он подчеркивает, что если бы зомби были возможны, то проблема сознание—тело могла бы оказаться нерешаемой. Это замечание помогает уточнить понимание Чалмерсом специфики проблемы сознания на раннем этапе.

Существо его позиции сводится к следующему. Во-первых, он убежден, что функционирование мозга может быть исчерпывающе объяснено с точки зрения третьего лица, методами когнитивных дисциплин и нейронауки. Иными словами, он не сомневается в каузальной замкнутости физического мира. Во-вторых, Чалмерс призывает всерьез относиться к субъективным данным от первого лица. Они являют собой своего рода «аномалию», но такую аномалию, которую нельзя игнорировать, а приходится объяснять. И для раннего Чалмерса проблема состоит прежде всего в том, как эти данные от первого лица вписываются в объективный порядок существования: распутав этот узел, мы, собственно, и объясним возможность «первого лица», о которой настойчиво вопрошает Чалмерс. Идея решения, которая уже вызрела у него к тому времени, подразумевала демонстрацию того, что субъективная реальность является обратной стороной объективных функциональных схем в мозге. Он выражает это отношение связкой «паттерн-информация». Трудно спорить, что объективные функциональные схемы в мозге, паттерны, являются носителями информации, и Чалмерс полагает, что информация, которую они несут, это и есть наши субъективные состояния.

Эта теория получает дальнейшее развитие в рукописи «Сознание и познание», созданной в январе 1990 г. Здесь Чалмерс поясняет, что предлагаемая им концепция является разновидностью так называемой «двуаспектной теории», вариации на тему которой он находит, в частности, в трудах Т. Нагеля. Речь идет о том, что качественные ментальные состояния — не что иное, как один из аспектов некоей реальности (за неимением лучшего Чалмерс называет ее «паттерн-информация»), другим аспектом которой являются функциональные схемы, паттерны. На первый взгляд неясно, что дает это уточнение. Кажется, что оно, скорее, запутывает ситуацию. Дальнейшие разъяснения Чалмерса, однако, помогают понять выгоды этого решения. Дело в том, что он характеризует принимаемый им вариант двуаспектной теории как вариацию на тему теории тождества — той самой теории Плейса и Смарта, которую мы уже разбирали. Теория тождества ментальных и

физических (или физически истолкованных функциональных) состояний, несмотря на множество трудностей, имеет одно важное преимущество: она позволяет изящно решить проблему ментальной каузальности. Здравый смысл говорит нам, что ментальные состояния играют определенную каузальную роль в поведении. Однако если мы признаем каузальную замкнутость физического (а в 1990 г. Чалмерс пока безоговорочно признает ее), то возникает проблема объяснения наших обыденных интуиций, касающихся ментальной каузальности. Теоретики тождества легко согласуют одно с другим: если ментальные состояния тождественны физическим процессам в мозге или физически реализованным функциональным схемам в нем, то мы имеем полное право сказать, что такие ментальные состояния действительно влияют на поведение.

Именно это обстоятельство привлекает Чалмерса: «субъективные состояния могут быть каузально действенными, так как они всего лишь другой аспект объективных состояний»⁶. Вместе с тем защищаемый Чалмерсом двуаспектный вариант теории тождества, уверяет он (хотя истинность его тезиса неочевидна), отличается от стандартных теорий признанием реальности квалиа: это «теория тождества, которая всерьез относится к субъективным состояниям».

Другое отличие своей теории Чалмерс видит в том, что она не ограничивается абстрактными рассуждениями о двух аспектах, а специфицирует их в уже упоминавшихся терминах «паттерна и информации». Паттерны существуют в мозге, а «квалиа — не что иное, как информация»⁷. Впрочем, в рукописи 1990 г. нет сколько-нибудь развернутого обсуждения понятия информации. Зато здесь присутствуют любопытные размышления о том, почему он считает свою теорию единственным перспективным подходом к решению проблемы сознание—тело. Размышления эти в зачаточном виде присутствовали и в рукописи «Точки зрения», но в разбираемом тексте они доведены прямо-таки до кипения. Чалмерс обращает внимание, что признаваемый им тезис о возможности исчерпывающего объяснения человеческого поведения с позиции третьего лица (являющийся следствием принципа каузальной замкнутости физического) неизбежно приводит к выводу, что с этой объективной позиции объяснимы и все высказывания, которые мы делаем о качественных состояниях и сознании. Получается, таким образом, что хотя «сознание таинственно, высказывания о сознании таковыми не являются».

Иными словами, все наши рассуждения о сознании и проблеме сознание—тело, будучи фактами поведения, могут быть

объяснены и без привлечения сознания (в конце статьи Чалмерс предлагает набросок такого объяснения, напоминающий схемы Деннета). Даже если бы мы были лишены качественных состояний, мы все равно говорили бы о них. Эта парадоксальная ситуация, четко зафиксированная в 1989 г. А. Элитцуром⁸ (на статью которого Чалмерс, кстати говоря, ссылается), требует какой-то реакции. Сам Элитцур отреагировал на нее отказом от тезиса о каузальной замкнутости физического и признанием реальной действительности ментальных состояний, нетождественных физическим процессам в мозге и ответственных за высказывания об этих состояниях. Чалмерс, однако, пока не готов поступиться каузальной замкнутостью физического.

Но как же тогда «сознание может иметь хоть какое-то отношение к тому, что мы говорим о сознании»? Двухаспектная теория отвечает на этот вопрос и позволяет сохранить «когерентность» точек зрения от первого и от третьего лица. Будучи оборотной стороной функциональных схем, отвечающих за высказывания о сознании, сознание действительно может иметь отношение к последним.

Таким образом, двухаспектная теория Чалмерса позволяет обосновать косвенную каузальную действительность (каузальную релевантность) качественных ментальных состояний. В этом ее принципиальное отличие от традиционного эпифеноменализма, к которому он объективно близок, как и всякий, кто признает онтологическую самостоятельность сознания но при этом допускает каузальную замкнутость физических процессов. Эпифеноменалисты тоже утверждают реальность качественных состояний и их скоррелированность с физическими процессами в мозге, но, не отождествляя их с определенными аспектами последних, они, по Чалмерсу, вынуждены говорить о случайном характере этой корреляции и не в состоянии объяснить, как она возможна. Именно поэтому в 1990 г. Чалмерс считает стандартный эпифеноменализм неприемлемой позицией. Вместе с тем он подчеркивает, что он перенимает от эпифеноменализма серьезное отношение к квалиа. Но он дополняет признание реальности качественных состояний их интерпретацией в качестве внутреннего аспекта функциональных паттернов в мозге. В итоге получается такой эпифеноменализм, в котором «сознание имеет значение» — но это все-таки эпифеноменализм, потому что «мы можем вообразить субъективные аспекты "оторванными" от несубъективных, допуская полную автономию физического».

Еще раз подчеркнем специфику отношения Чалмерса к эпифеноменализму в 1990 г. В классическом виде, полагал Чалмерс, эпифеноменализм неприемлем, так как не позволя-

ет объяснить корреляцию квалиа и физических процессов в мозге — ключевое слово здесь «объяснить». Двухаспектная теория дает такое объяснение и поэтому выглядит более перспективной. Мы видим, таким образом, что теория Чалмерса и в самом деле позволяет решать какие-то проблемы. С другой стороны, нельзя не отметить, что попытки Чалмерса конкретизировать упомянутую двухаспектность не только решают проблемы, но и порождают немало проблем. В частности, истолкование качественных состояний как информации, носителями которой являются объективные паттерны, подталкивает к смелым выводам о широте распространенности сознания в мире. В самом деле, паттерны и функциональные схемы реализованы не только в мозге. Информация и ее носители, собственно, есть повсюду. Значит ли это, что сознание — как внутренний аспект паттернов — тоже имеет повсеместное распространение? Характерно, что в рукописи 1990 г. Чалмерс без особых колебаний соглашается с этим предположением. Он, правда, уточняет, что в таких случаях речь надо вести, скорее, не о сознании как таковом — предполагающем наличие сложных механизмов самомониторинга, — а о квалиа, но вот они-то вполне могут существовать даже в примитивных системах обработки информации, к которым можно причислить вообще все, что только пожелаешь.

Симпатия к панпсихизму в дальнейшем сослужила неоднозначную службу Чалмерсу. Хотя она, возможно, способствовала распространению его взглядов в нефилософской среде, она вызывала резкую критику со стороны ряда академических философов. Сейчас, однако, важно понять, почему Чалмерс не побоялся панпсихистских выводов, столь очевидно расходящихся со здравым смыслом. Ответ в том, что эти выводы являются следствием его тезиса о неразрывной связи объективных паттернов и субъективной информации.

Но почему Чалмерсу так важно было говорить именно об их неразрывной связи? Вспомним его загадочное замечание из незавершенной статьи 1989 г. о том, что, если бы зомби были возможны, проблема сознание—тело оказалась бы неразрешимой. По-видимому, он считал, что допущение возможности зомби означает отрицание неразрывной связи данных от первого и третьего лица. Соответственно, он был уверен, что решение проблемы сознание—тело предполагает демонстрацию такой связи. Более того, первоначально Чалмерс, скорее всего, полагал, что его двухаспектная теория позволяет проиллюстрировать связь такого рода. Однако в 1990 г. он начинает сомневаться, а так ли это в действительности? В самом деле, даже если признать, что качественный субъективный опыт есть внутренний аспект неких объективных процессов, все

равно остается вопрос, почему этот аспект так уж необходим. Не случайно, что теперь он более определенно допускает представимость зомби. Осмысление им факта незавершенности своей теории сопровождалось примечательным изменением самой формулировки проблемы сознания.

Об этом изменении следует сказать особо. В рукописи 1989 г. Чалмерс рассуждает о тайне сознания и данных от первого лица в контексте «кантианского» вопроса «как возможно?» Специфика этого вопроса в том, что, задавая его, мы уже предполагаем данность квалиа и пытаемся понять, каким образом можно встроить эти «аномальные» состояния в объективный порядок вещей. Двухаспектная теория помогает справиться с этой задачей. Однако даже она не позволяет продемонстрировать необходимую связь субъективной и объективной реальностей. Что-то остается необъясненным. В рукописи 1990 г. Чалмерс нащупывает формулировку, которая выявит пробелы, имеющиеся в его теории, и будет стимулировать его к дальнейшим поискам. Эта формулировка предполагает дополнение вопроса «Как возможны данные от первого лица?» вопросом «Почему такие субъективные состояния вообще должны существовать?». Вариации на тему «Почему существуют квалиа?» буквально наводняют текст 1990 г. и полностью отсутствуют в более раннем наброске.

Рождение нового вопроса произошло, по-видимому, не без влияния статьи Ф. Джексона «Эпифеноменальные квалиа», на которую Чалмерс ссылается в рукописи 1990 г. «Эпифеноменальные квалиа, — писал Джексон, — не имеют никакого отношения к выживанию. Ни на каком этапе нашей эволюции естественный отбор не благоприятствовал тем, кто мог бы прояснить, как они порождаются и каким законам подчиняются, — да и почему они вообще существуют»⁹. Финальный пассаж содержит новый вопрос Чалмерса.

Нельзя, конечно, утверждать, что вопросы «почему существуют квалиа?» и «как возможны квалиа?» никак не связаны. Сам Чалмерс сближает их, замечая, что ответ на первый вопрос подразумевает объяснение того, «как возможен субъективный опыт». Но из этого не следует, что объяснение того, «как возможен субъективный опыт», подразумевает ответ на вопрос «почему мы обладаем сознанием?». Ведь некоторые объяснения, как мы только что убедились, оставляют этот вопрос без ответа: двухаспектная теория позволяет понять возможность корреляции данных от третьего лица с данными от первого лица (она возможна именно потому, что это два аспекта одной и той же реальности), но не помогает понять, почему данные от первого лица вообще существуют. Таким образом, новая акцентировка проблемы сознания позволяет выявить

серьезную лакуну чалмерсовской двуаспектной теории. Неудивительно, что в конце статьи Чалмерс признает, что представленный в ней набросок теории пока не позволяет разрешить проблему сознание—тело. Неудивительно и то, что Чалмерс объявляет вопрос о том, почему существуют квалиа, «глубочайшей» из всех проблем, связанных с сознанием.

В скором времени Чалмерс нашел путь преодоления этих трудностей и на рубеже 1992 и 1993 гг. за шесть месяцев создал первый вариант трактата «Сознающий ум: в поисках фундаментальной теории», в котором намечена одна из версий решения «глубочайшей» проблемы. Впрочем, прежде чем выйти в свет в 1996 г., «Сознающий ум», изначальный вариант которого предстал в виде докторской диссертации Чалмерса под названием «На пути к теории сознания» (1993)¹⁰, претерпел немало доработок: в версии 1993 г., к примеру, отсутствовали (хотя и анонсировались) последние главы «Сознающего ума», в которых Чалмерс как раз и предлагает свое решение упомянутой проблемы. В целом диссертация ориентирована скорее на двуаспектную теорию (рассматривающую вопрос, как возможны квалиа, т. е. как они встроены в физическую реальность), чем на обсуждение «глубочайшей проблемы». В «Сознающем уме» картина иная. Но еще до его публикации Чалмерс, как мы знаем, поделился с научным сообществом своими соображениями по поводу найденной им новой формулировки проблемы сознания и решения последней. Он сделал это и в устной форме в Туссоне в 1994 г., и в письменной — в статье «Навстречу проблеме сознания», первоначально опубликованной в *Journal of Consciousness Studies* в 1995 г.

Здесь «глубочайший» вопрос о том, почему вообще существуют субъективные состояния, — который, напомним, именно в такой формулировке имеет мало аналогов в прошлом¹¹ (хотя это, разумеется, не отрицает его связи с традиционными интерпретациями проблемы сознание—тело), -наконец-то появляется под уже хорошо известным нам именем «трудной проблемы сознания». В диссертации 1993 г. этот термин отсутствовал, хотя уже там было обозначено противопоставление «трудной» и «легкой» части проблемы сознание—тело¹². Вместе с тем пассаж из введения в «Сознающий ум» (в целом совпадающего с введением в диссертацию), где Чалмерс использует выражение «трудная проблема», -позднейшая вставка, и в диссертации этой фразы нет. Одним словом, хотя в диссертации Чалмерс уже говорит о «трудном» и «легком» аспектах проблемы сознания, он еще не употребляет выражения «трудная проблема»¹³. Когда же произошло рождение этого знаменательного термина? Вот что сообщает

нам сам Чалмерс: «Вспоминаю, что использовал различие трудной и легких проблем на моем семинаре по сознанию в Вашингтонском университете в конце 1993 г. Я не припоминаю, чтобы я пользовался этой фразой до этого. Публично я впервые употребил это выражение на Туссанской конференции в апреле 1994 г. Я и не думал, что оно произведет такой эффект. Поэтому я превратил свое выступление в статью и вставил дополнительный абзац во введение к книге»¹⁴. Статья, о которой говорит Чалмерс, это, собственно, и есть упомянутая выше статья «Навстречу проблеме сознания».

Начинается она как раз с различия «трудной» и «легких» проблем. Прежде всего, утверждает Чалмерс, надо выделить из всего пула проблем сознания «легкие» проблемы, а именно те проблемы, которые можно решить «стандартными методами когнитивной науки, объясняющей тот или иной феномен с помощью вычислительных или нейронных механизмов»¹⁵. К числу легких проблем он относит те, что имеют отношение к объяснению феноменов «различения, категоризации», «интегрирования информации», «способности выдавать отчеты о ментальных состояниях», способности системы к самомониторингу, а также проблемы внимания, контроля за поведением и объяснения различий между сном и бодрствованием.

Все перечисленные проблемы (список которых в этой статье, разумеется, не претендует на полноту), как подчеркивает Чалмерс, соотнесены с проблемой сознания, но последняя не исчерпывается ими. Даже если мы установили физические механизмы, отвечающие за упомянутые выше процессы, все равно остается вопрос, выражающий «действительно трудную проблему сознания»: почему с этими процессами связан субъективный опыт (experience)? «Многие согласны, — пишет Чалмерс, — что опыт возникает на физической основе, но мы не можем как следует объяснить, почему и как он возникает подобным образом»¹⁶. Иными словами, «почему вся эта обработка информации [в мозге] не происходит "в темноте", безо всякого внутреннего чувства? Почему, когда электромагнитные волны различной частоты воздействуют на сетчатку, различаются и категоризируются визуальной системой, это различие и категоризация испытываются как ощущение чего-то ярко-красного? Мы знаем, что при выполнении этих функций возникает сознательный опыт, но главная тайна как раз и состоит в самом факте его возникновения»¹⁷.

Обозначив «трудную проблему», Чалмерс показывает, что большинство авторов, изучающих сознание, уходит от нее. Под огонь попадают и Крик с Кохом, и Баарс с Эдельманом.

Все они занимаются отысканием нейронных или функциональных коррелятов сознания, но даже при условии успешности их проектов вопрос, почему те или иные нейронные или вычислительные процессы сопровождаются субъективным опытом, остается у них без ответа. Но было бы странно, если бы без ответа оставил его и сам Чалмерс. И он действительно намечает подход к решению «трудной проблемы».

Суть этого подхода состоит в признании фундаментальности сознания или «опыта» и допущении существования столь же фундаментальных законов, связывающих сознание с физическим миром. Сразу возникает сомнение, является ли это решением «трудной проблемы». Можно ли отвечать на вопрос «Почему существует сознательный опыт?» утверждением о том, что он существует как некая нередуцируемая реальность? Отчасти это, конечно, ответ: сознание существует потому, что оно фундаментально, т. е. не может не существовать. Но тогда можно спросить, почему сознание фундаментально, - и мы не сдвинулись с места. Чалмерс предвидит это возражение: «Конечно, если мы допускаем фундаментальность опыта, то в некотором смысле этот подход не отвечает нам на вопрос, почему опыт вообще существует». «Но, — продолжает он, — это справедливо для всякой фундаментальной теории. Ничто в физике не говорит нам, почему вообще существует материя, но мы не протестуем против теорий материи. Любая научная теория должна признавать фундаментальность определенных характеристик мира»¹⁸. С этим замечанием трудно спорить, но тогда возникает вопрос, а стоило ли Чалмерсу в подобном случае так акцентировать «трудную проблему»? В конце концов, не говорим ведь мы о трудной проблеме материи и массы! Загвоздка и в том, что при постановке «трудной проблемы» он уже исходил из нередуцируемости сознания, и непонятно, как эта проблема может быть решена повторением условий ее постановки.

Один из вариантов интерпретации — признать, что весь чалмерсовский дискурс «трудной проблемы» имеет герменевтический характер и предназначен исключительно для прояснения и верификации обыденной интуиции о несводимости сознательных состояний к физическим процессам в мозге. Но этот вариант плохо сочетается с изначальными интенциями Чалмерса, известными нам из его ранних рукописей. Мы видели, что он всегда хотел показать неразрывную связь между сознанием и мозгом, но такую связь, которая не отменяла бы онтологических различий между ними. Иначе говоря, в конечном счете он стремился к объяснению, почему определенные процессы в мозге сопровождаются сознательным опытом.

Таким образом, можно предположить, что тезис Чалмерса о фундаментальности сознательного опыта все же может рассматриваться в качестве объяснения того, почему он вообще существует. Это предположение можно подтвердить на материале другой статьи Чалмерса — «Продвигаясь к проблеме сознания», опубликованной в 1997 г.¹⁹ в том же журнале, что и «Навстречу проблеме сознания», и содержащей ответ на 26 отобранных редакцией комментариев на ту статью 1995 г. (в 1997 г. обе статьи и комментарии были изданы и отдельной книгой: «Объясняя сознание: трудная проблема»).

В статье 1997 г. Чалмерс рассуждает о том, что такое объяснение вообще, — и приводит пример с Ньютоном²⁰. Если бы Ньютон, осмысляя законы движения, ограничился констатацией каких-то внешних регулярностей, то это еще нельзя было бы назвать объяснением. Но за подобными регулярностями он усмотрел общий закон тяготения. И, ссылаясь на этот закон, он уже мог объяснять конкретные события, такие, к примеру, как падение яблока ему на голову. Т. е. если мы спросим, почему падает яблоко, можно ответить, что оно падает, потому что подчиняется закону всемирного тяготения. Точно так же, если спросить, почему те или иные процессы в мозге сопровождаются сознанием, то Чалмерс ответит нам, что это происходит в силу некоего фундаментального психофизического закона.

Но в чем состоит этот психофизический закон? Прежде чем найти его, замечает Чалмерс, необходимо оговорить рамки, которым он должен соответствовать. Эти рамки могут быть заданы более поверхностными, чем предполагаемый базовый психофизический закон, «высокоуровневыми» принципами корреляции ментального и физического, которые, по мнению Чалмерса, легче обнаружить и обосновать. В статье 1995 г. он называет два таких принципа: «структурной когерентности» и «организационной инвариантности». Первый из них эксплицирует систематическую связь между «сознанием» (consciousness) и «осведомленностью» (awareness), второй констатирует, что факт существования сознательного опыта и характер его наполненности зависят исключительно от функциональной организации связанной с ним физической системы.

Чтобы понять, почему Чалмерс останавливается именно на этих принципах, еще раз напомним контекст. Он хочет найти закон, связывающий физические процессы и качественные ментальные состояния. Этот закон позволит предсказывать существование подобных состояний на основе знания о физических кондициях той или иной материальной системы. Соответственно, поиск подобного закона лучше всего начинать с системы, качественная начинка которой доподлин-

но известна исследователю, а именно с нашего собственного организма. Что же можно сказать по этому поводу? Прежде всего, концептуальный анализ наших осознанных состояний показывает (точнее, мог бы показывать — сам я думаю, что это не так, но сейчас это неважно), что если я осознаю нечто, то это нечто как данность сознания не только может учитываться мной при планировании моих действий, но и, как кажется, оказывает прямое влияние на мое невербальное и вербальное поведение. В частности, если я осознаю нечто, обычно я могу отчитаться о наличии этой данности.

Исходя из каузальной замкнутости физического и скоррелированности физических процессов с качественными состояниями, логично предположить, что сказанное можно выразить и на языке функций материальных систем. Для этого Чалмерсу и потребовался термин «осведомленность»), коррелятивный «сознанию», но лишенный качественного оттенка. Материальная система осведомлена в чем-то, когда это нечто непосредственно доступно для тех ее подсистем, которые отвечают за формирование ее поведения. Говоря словами Чалмерса, осведомленность есть «непосредственная доступность для глобального контроля». Если принять этот принцип - очевидно перекликающийся с баарсовской моделью «глобального рабочего пространства»²¹, — то кандидатами на наличие сознательных состояний будут системы, где происходит модульная обработка информации, поступающей затем в глобальное рабочее пространство, где она используется для подготовки и совершения действий.

Принцип структурной когерентности не ограничивается, однако, указанием необходимых критериев наличия сознательных состояний. Он позволяет узнать нечто и об их внутренней организации. Ведь структура сознательного опыта отражает особенности протекания объективных когнитивных процессов, ответственных за его появление. Скажем, если на качественном уровне можно различать цвета, то это различие должно существовать и на физическом уровне, представляя в виде активации разных групп нейронов и т. п. Кстати, в диссертации 1993 г., а также в соответствующих главах «Сознающего ума» Чалмерс объединяет под рубрикой принципа структурной когерентности именно такие структурные соответствия, выделяя тезис о соответствии сознания и осведомленности в особый принцип²². Но в статьях 1995 и 1997 гг. он не проводит этой дифференциации, что вполне понятно: принцип соответствия сознания и осведомленности тоже имеет структурный характер.

Тема структурного или организационного соответствия физических и качественных процессов и состояний получа-

ет продолжение и развитие при обсуждении Чалмерсом принципа организационной инвариантности. Важность этого принципа может быть проиллюстрирована следующим образом. Допустим, мы признали, что всякое осознанное состояние сопровождается определенной интеграцией тех или иных данных на физическом уровне. Но из этого не следует, что всякая подобная интеграция сопровождается сознанием. Принцип структурной когерентности может претендовать лишь на указание необходимых, но не достаточных критериев. Принцип организационной инвариантности, напротив, должен способствовать определению достаточных критериев сознания. Согласно этому принципу, за наличие в той или иной физической системе сознания отвечают именно ее функциональные компоненты. Т. е. если материальная система интегрирует поступающие в нее данные по такой же схеме, что и мой, к примеру, мозг, то такая система с необходимостью должна обладать теми же качественными состояниями, что и я.

Нарисованная картина соотношения принципов когерентности и организационной инвариантности является, впрочем, несколько упрощенной. Дело в том, что в статье 1995 г. Чалмерс выражается так, будто принцип когерентности позволяет задавать не только необходимые, но и достаточные условия наличия сознания²³. Однако в «Сознающем уме» Чалмерс проясняет ситуацию²⁴, и его уточнения позволяют нам лучше понять диалектику двух принципов. Говоря о принципе когерентности, он отмечает, что вопрос о «достаточности осведомленности» для сознания, в отличие от вопроса о его необходимости для него, «является более трудным»²⁵. Ведь можно допустить, что для порождения сознания материальная система должна не только быть осведомлена в чем-то, но и обладать каким-то дополнительным свойством, неким «X-фактором»²⁶. Чалмерс, правда, утверждает, что допущение подобного фактора в любом случае будет чем-то произвольным, а сам он — случайным. Иными словами, тезис, что сознание возникает при наличии осведомленности и еще какого-то X-фактора, выглядит, с его точки зрения, явным отступлением от требований простоты научных теорий — тем более что хорошего кандидата на роль X-фактора найти не удастся. В пятой главе я покажу, что это заключение сомнительно. Но важно, что и сам Чалмерс понимает, что он привел не очень сильный довод²⁷: требование простоты не имеет обязательного характера, да и является внешним по отношению к принципу когерентности.

Поэтому ему нужны дополнительные аргументы, которые, в частности, показывали бы достаточность осведомленности для сознания. Принцип, формулируемый на основе этих аргу-

ментов, будет отличаться от принципа структурной когерентности тем, что в его формулировке будет выражена обусловленность сознания функциональными схемами, реализованными в материальных системах. Речь, конечно, идет о принципе организационной инвариантности, и теперь нам пора сосредоточиться на самих аргументах Чалмерса.

Главный из них изложен в статье 1995 г., но в полном объеме они развернуты в других текстах — статье «Отсутствующие квалиа, блекнувшие квалиа, скачущие квалиа», написанной в 1993 г.²⁸, и в «Сознающем уме». Чалмерс предлагает два мысленных эксперимента, обладающих, по его мнению, большим демонстративным потенциалом. Первый из них должен показать реальную невозможность того, что реализация функциональных схем, которые сопровождаются сознанием в моем случае, на других носителях может существовать и при отсутствии сознания, т. е. при отсутствующих квалиа. Второй показывает, что в подобном случае квалиа не только не могут отсутствовать, но и не могут отличаться от тех, что есть у меня²⁹.

Первый мысленный эксперимент разворачивается следующим образом: допустим, мой функциональный изоморф, сделанный, к примеру, из кремниевых чипов, не обладает сознанием. Далее, допустим, что нейронаука и нейрохирургия достигли такого уровня, что мы можем заменить часть моих нейронных сетей на кремниевые чипы, полностью дублирующие функции замененных нейронов. При полной замене я стану точной копией изоморфа, который, по предположению, не обладает сознанием. Не буду, выходит, обладать им и я. Но что будет происходить в процессе замены, пока она лишь частичка? Очевидно, отвечает Чалмерс, мое сознание должно будет постепенно блекнуть. Парадокс, однако, в том, что при этом мое поведение, опять же по предположению, будет оставаться неизменным. Я буду, к примеру, говорить, что вижу яркие цвета, хотя в действительности от них останутся лишь их бледные подобию. И я буду говорить, что различаю их оттенки, субъективно утратившие для меня реальность. Такая ситуация, полагает Чалмерс, маловероятна, и это позволяет заключить, что замена нейронов на чипы не приведет к утрате сознания. Соответственно, для сознания важен не носитель, а функциональная организация определенной материальной системы.

Второй аргумент, который Чалмерс считал еще более сильным³⁰, выстраивается по аналогии с первым. Допустим, что мой функциональный изоморф наделен перевернутыми квалиа: там, где я вижу красное, он видит синее, и т. п. Опять проведем частичную замену моих нейронных сетей на крем-

ниевые дубликаты. После этой замены я буду воспринимать некий промежуточный цвет. Но не будем полностью выводить из игры старые нейронные цепи, а устроим все так, чтобы можно было вновь подключить их, — и подключим. Что произойдет? Ощущаемый цвет мгновенно поменяется — он опять станет красным. Теперь опять активируем кремниевый участок цепи — и цвет снова станет ближе к синему, это и есть скачущие квалиа. Но парадокс опять-таки в том, что несмотря на все эти скачки и танцы, я снова ничего не замечу — ведь кремниевая цепь функциональная изоморфна нейронной. Такая ситуация, по Чалмерсу, совершенно неправдоподобна, и это означает, что квалиа не только не должны исчезать (данный аргумент может доказывать и это), но и не должны меняться при реализации одинаковых функциональных схем на разных носителях. И если так, то мы показали, что наличие определенных функциональных схем является достаточным условием наличия сознания и достаточным условием для заключения о его характере.

Чалмерс подчеркивает, что данные мысленные эксперименты не могут показать и не показывают логической невозможности описанных ситуаций, но они и не предназначены для этого. Они призваны продемонстрировать реальную или «естественную» (natural) невозможность изменения или отсутствия квалиа в моем функциональном изоморфе, равно как и в системах, функционирующих подобно мне — можно преобразовать эти эксперименты так, чтобы они достигали именно такой цели. И если это так, то мы и в самом деле с уверенностью можем говорить, что все люди воспринимают мир примерно так, как я: наши различия возрастают сообразно отдалению реализованных у них функциональных схем от моей собственной.

Все, однако, не так радужно, как представляет Чалмерс. Проблема в том, что в описанных им случаях нет никаких парадоксов. Особенно это очевидно со скачущими квалиа. Попробуем еще раз встать на место человека, которому подключают то кремниевые чипы, то нейронные цепи. Допустим также, что при этом его квалиа действительно скачут. Что нас смущает? Что он будет вести себя так же, как он вел бы себя, если бы они не скакали? Что он не замечает этих скачков? Но поводов для смущения и удивления нет. Чтобы заметить отличие, надо сопоставлять наличное ощущение с тем, которое было раньше и, согласно предположению, отличалось от наличного. Но более чем вероятно, что нейронные или кремниевые цепи, участвующие в обработке зрительных данных, отвечают и за воспоминания о прежних случаях восприятия соответствующих цветов³¹. Замещая нейронную

цепь кремниевой, мы переписываем не только наличные ощущения, но и воспоминания о них. Таким образом, в подобной ситуации человек в принципе не может заметить какое-то отличие воспринимаемого им цвета от того, что воспринимался им раньше, — и действительно, по предположению, не замечает его.

Сходное возражение против аргумента скачущих квалиа выдвинул С. Шумейкер³². Чалмерс попытался ответить ему в статье «Материализм и метафизика модальности»³³ но, по правде говоря, его ответ не выглядит убедительным, поскольку Чалмерс продолжает настаивать на странности того, что субъект не будет замечать изменений, тогда как суть данного возражения как раз в том, что это вовсе не является странным. Тем не менее в той же статье Чалмерс сместил центр тяжести в сторону аргумента с блекнущими квалиа и заявил, что теперь именно этот аргумент он считает «более сильным»³⁴. Однако и этот аргумент, и мысленный эксперимент можно нейтрализовать, хотя его критику лучше выстраивать иным способом. Напомним, что видимость парадокса в этом эксперименте возникает потому, что человек с чипами, частично замещающими нейроны, утверждает, что фиксирует какие-то различия, к примеру, между оттенками синего цвета, в то время как на уровне осознаваемых им поблекших данностей эти различия становятся настолько малыми, что по идее не должны были бы констатироваться им. Но на это можно заметить, что, во-первых, эти различия все же не исчезают и на фоне отсутствия более ярких впечатлений они могут быть замечены³⁵. Во-вторых, наличие у него убеждения в их существовании, подкрепленного вербальными и поведенческими диспозициями, заставит его видеть различия даже там, где увидеть их, казалось бы, невозможно. И это самая обычная ситуация: концепты, задействованные в конструировании восприятия, как известно, могут творить чудеса и посильнее этого — благодаря им мы дорисовываем контуры несуществующих фигур и часто видим различия даже там, где их нет.

В общем, приходится признать, что аргументы Чалмерса, приведенные им в поддержку принципа организационной инвариантности в перечисленных выше статьях и в седьмой главе «Сознающего ума», посвященной именно этому принципу, не достигают цели. Эти принципы могут быть истинными, но это надо еще доказать. К данной теме мы обязательно вернемся, сейчас же следует напомнить, что чалмерсовские принципы предполагают каузальную замкнутость физических процессов. В самом деле, если допустить, что качественные состояния оказывают реальное влияние на поведение, то эти состояния должны быть встроены в функциональные схемы,

определяющие наши реакции на те или иные внешние импульсы. Но тогда невозможно будет говорить, что наше сознание порождается реализацией таких функциональных схем в мозге или зависит от них.

Понимание того обстоятельства, что рассуждения Чалмерса о принципах структурной когерентности и организационной инвариантности ведутся им на основе допущения каузальной замкнутости физического, возвращает нас к проблеме эпифеноменализма. Напомню, что в рукописи 1990 г. «Сознание и познание» Чалмерс утверждал, что эпифеноменализм, отрицающий каузальную действенность качественных состояний в физическом мире, — несостоятельная философская доктрина, и обещал, что его теория соотношения ментального и физического позволит преодолеть ее недостатки. Принципы структурной когерентности и организационной инвариантности являются одним из компонентов его новой теории. Исходя из этого, можно предположить, что Чалмерс рассматривает их в антиэпифеноменалистском контексте. С другой стороны, поскольку эти принципы, как и эпифеноменализм, опираются на тезис о каузальной замкнутости физического, данное предположение нуждается в текстуальном подтверждении. Текстуальное подтверждение в этой связи и правда можно найти, но оно подтверждает совсем другое.

А именно, Чалмерс заявляет, что принципы структурной когерентности и организационной инвариантности как раз совместимы с эпифеноменализмом. Он рассуждает об этом в статье «Сознание и его место в природе», написанной в 2001 г. и впервые опубликованной в сборнике «Философия сознания: классические и современные тексты» (2002). Эпифеноменалист, отмечает Чалмерс, признает «существование фундаментальных психофизических законов, связывающих физические и феноменальные свойства»³⁶. Любопытно и то, что и в этом, и в других текстах, а именно в «Сознающем уме» и статье «Продвигаясь к проблеме сознания» Чалмерс больше не называет эпифеноменализм несостоятельной философией³⁷. Наоборот, он подчеркивает, что эпифеноменализм лишен «фатальных дефектов», а в «Сознающем уме» идет еще дальше и не исключает, что он может оказаться единственной приемлемой позицией³⁸.

Потепление чалмерсовского отношения к эпифеноменализму, впрочем, вовсе не означает, что он идентифицирует свои взгляды как эпифеноменализм. Скорее, оно является выражением определенной неуверенности, имеющейся у него по поводу своей позиции. Тем не менее позиция эта у него есть, и он вовсе отказался — по крайней мере если речь идет о середине 90-х гг. — от намерения преодолеть «неэлегант-

ность» эпифеноменализма (согласно которому качественное оказывается неким довеском к физическому).

Кроме того, Чалмерс не забыл и о своих прежних упреках, высказанных в адрес этой концепции: эпифеноменализм так или иначе оставляет ощущение случайности качественных состояний, ощущение того, что их могло бы и не быть: «трудная проблема сознания» остается по большому счету нерешенной. Но теперь Чалмерс полагает, что, возможно, с этим придется смириться, просто постулировав те или иные принципы, наводящие мосты между физическим и ментальным, и удовлетворившись «объяснениями» существования сознания через порождающие его «силы» или законы. Однако это все же не последнее слово Чалмерса. В конце концов, принципы структурной когерентности и организационной инвариантности, инкорпорированные эпифеноменализмом, не являются, как мы знаем, базовыми психофизическими законами. Возможно, они есть лишь частные случаи более фундаментального закона. Возможно также, что анализ этого закона откроет перспективы преодоления эпифеноменализма или хотя бы его модификации в духе двуаспектной теории, на которую Чалмерс возлагал большие надежды в начале 90-х.

Чалмерс действительно работал в этом направлении. Результаты его поисков в полном объеме изложены в «Сознающем уме», упоминаются они и в статье «Навстречу проблеме сознания» и др. Представленная в «Сознающем уме» концепция и в самом деле является наследницей двуаспектной теории из рукописи 1990 г. Ключевую роль в ней, напомним, играло понятие информации. Чалмерс утверждал, что качественные состояния и есть информация, носителями которой оказываются функциональные схемы (patterns) в мозге. Он также утверждал, что информация и схемы, или паттерны, — это два аспекта одной и той же вещи (для именованной которой он использовал мало что разъясняющий термин «паттерн / информация»), и поэтому с некоторыми оговорками причислял себя к теоретикам тождества. Но в «Сознающем уме» речь об этом уже не заходит.

Кроме того, в позиции Чалмерса произошли следующие изменения. Прежде всего, он основательно трансформировал и конкретизировал свое понимание информации. Опираясь на идеи К. Шеннона, он дефинирует ее как систему различий и отношений в абстрактном логическом пространстве³⁹. Информация имеется везде, где осуществлен выбор из некоего спектра возможностей. Сама по себе информация абстрактна, но она может быть реализована. Наиболее наглядной является физическая реализация. К примеру, у нас есть информационное пространство, где состояние А соотнесено с состоянием

В так, что отсутствие А влечет за собой отсутствие В. Физической реализацией этого пространства может быть, скажем, выключатель света. Это устройство обладает рядом физических свойств и может оказаться в разных состояниях, но если смотреть на него с точки зрения реализации определенной информационной матрицы, то эти состояния сводятся к двум: «включен» и «выключен» — разного рода физический «шум» отсекается потому, что содержащиеся в нем различия не имеют значения для данной системы, не производят в ней различий: согласно известной формуле Г. Бейтсона, информация — «это различие, которое производит различие»⁴⁰.

Итак, информация может быть реализована физически, и материальная система, реализующая ее, функционирует в соответствии с логическими отношениями, конституирующими информацию. Но тут возникает вопрос: а возможны ли иные способы реализации информации? Чалмерс уверен в позитивном ответе. Информация, полагает он, может быть реализована и в нашем качественном опыте. Ведь и в нем присутствует множество различий и соотношений.

Понятие информации, раскрываемое в терминах различий и соотношений в логическом пространстве возможностей, несомненно выглядит более фундаментальным, чем, скажем, понятие осведомленности. И поскольку Чалмерс заинтересован в отыскании фундаментальных психофизических законов, неудивительно, что он связывает их поиски именно с понятием информации. Преимущество этого понятия еще и в том, что высокоуровневые принципы структурной когерентности и организационной инвариантности легко могут быть истолкованы в информационных терминах. Так, согласно принципу структурной когерентности, различия, фиксируемые в сознательном опыте, отражаются в объективных различиях связанных с ним материальных систем — это наглядный пример двоякой реализации одной и той же информации. Принцип организационной инвариантности также допускает информационную интерпретацию. Дело в том, что всякая функциональная организация может быть представлена в качестве специфически определенного информационного пространства, реализованного на физическом уровне — и изоморфно отображенного в качественном опыте.

Помимо возможности наведения мостов с высокоуровневыми психофизическими принципами новая концептуальная схема Чалмерса может индуцировать плодотворные онтологические решения. Еще раз подчеркнем, что если в 1990 г. Чалмерс утверждал, что квалиа — это и есть информация, то теперь он говорит о квалиа как об одном из способов реализации информации. Такая трактовка открывает возможность создания бо-

лее стройной двуаспектной теории. Загадочная вещь, именуемая им ранее «паттерном/информацией», получает название «информация», и эта единая информация допускает физическую, т. е. внешнюю, и качественную, внутреннюю реализацию. Качественный «опыт, — пишет Чалмерс, — это информация изнутри, физика — информация извне»⁴¹.

Кому-то может показаться, что в «Сознающем уме» Чалмерс просто сменил терминологию. Но это не так. В рукописи 1990 г. он трактует «паттерн-информацию» как реальную вещь, тогда как в трактате 1996 г. он дает понять, что информация как точка схождения качественного и физического — это, скорее, абстрактный фокус. Именно поэтому Чалмерс больше не позиционирует себя как сторонника теории тождества. Но почему же он не развивает заманчивую идею гипостазирования информации и трактовки качественного и физического как двух эманативных аспектов последней? Дело, видимо, в том, что такая позиция не позволяла бы продвинуться в решении «трудной проблемы», понять, почему вообще существует качественный аспект информации. Чалмерс, судя по всему, осознал тупиковость этого пути еще в начале 90-х. Примерно в то же время он наткнулся — не без помощи американского философа Г. Розенберга⁴² - на другой вариант онтологии, который, собственно, и стал для него приоритетным. Его основу составляют идеи Б. Рассела, высказанные им в книге «Анализ материи», опубликованной еще в 1927 г. Рассел полагал, что имеющиеся у нас понятия о материальных свойствах имеют структурный характер, т. е. де-финируются через отношения чего-то одного к чему-то другому. Иными словами, физический мир предстает нам как некая система отношений — но без представления о том, что, собственно, вступает в отношения. Однако это что-то, несомненно, должно существовать. Таким образом, у материи должны иметься не только внешние, относительные, но и какие-то внутренние свойства. Что же представляют собой эти свойства? Рассел не исключал, что они могут иметь ментальный характер⁴³, — и это резонно: других внутренних свойств мы просто не знаем.

Эта идея — именно то, что было нужно Чалмерсу. Теперь он наконец получил возможность решить «трудную проблему», дать ответ на вопрос о том, почему помимо физических процессов в мозге существует и связанный с ними внутренний опыт. Он существует потому, что без него не существовала бы и физическая реальность мозга, опирающаяся на квалиа как на свой фундамент. Двуаспектная теория Чалмерса получает окончательное завершение. Два аспекта информации, внешний и внутренний, неразрывно связаны.

Самое интересное, однако, то, что в «Сознающем уме» мы не слышим победных реляций по поводу решения «трудной проблемы». Чалмерс очень осторожно рассуждает на этот счет. Неорасселовская теория ментального как основы физического подается им как спекулятивная гипотеза, красивая и элегантная, позволяющая построить монистическую картину мира (дуализм предполагает равноправие ментального и физического, а здесь его нет), — но не более.

Осторожность Чалмерса объясняется целым рядом причин. Во-первых, для него не вполне очевидно, что рассуждения Рассела относительно несамодостаточности физических аспектов материи верны, — хотя он и склоняется к такому выводу. Во-вторых, он осознает, что признание истинности расселовской модели может привести к панпсихизму. В самом деле, если ментальное является онтологической предпосылкой физического, то оно должно быть повсюду. Нельзя сказать, что Чалмерс опасается объявлять себя панпсихистом: он специально показывает, что в панпсихизме нет явных нестыковок. Кроме того, панпсихизм хорошо сочетается с глобальным информационным подходом Чалмерса и с его трактовкой качественных состояний как одного из способов реализации информации. Вместе с тем он, конечно, соглашается, что говорить о богатой феноменальной жизни, скажем, термостата или тем более электрона нелепо. Внутренние свойства подобных систем, если они вообще существуют, лишь очень отдаленно напоминают состояния нашего сознания. И это приводит его к предположению, что внутренние состояния материи имеют, скорее, протофеноменальный, «протоопытный», а не собственно качественный ментальный характер.

Эту модификацию расселовской теории Чалмерс именуется «панпротопсихизмом» — и именно она вызывает у него наибольшую симпатию. Однако даже панпротопсихизм не устраивает его в полной мере. Проблема прежде всего в том, что неясно, как понимать «протоопытные» свойства. Насколько они вообще известны нам? Если сказать, что они известны нам в том плане, что лишь несущественно отличаются от привычных нам качественных состояний, то тогда придется признавать, что физическое немислимо без ментального. А из этого, в частности, следует непредставимость зомби. Мы помним, правда, что в ранних рукописях Чалмерс действительно надеялся показать, что зомби непредставимы. Но трудность в том, что в самом «Сознающем уме» Чалмерс использует тезис о представимости зомби в качестве одного из главных доводов в пользу нередуцируемости качественных состояний к физическим свойствам: если зомби представимы, то квалиа как бы дополнительные к ним.

Как выйти из этой ситуации без противоречий? В «Сознающем уме» и в других текстах, в частности в статьях «Сознание и его место в природе» и «Двумерный аргумент против материализма», Чалмерс разрабатывает для этого аппарат «двумерной семантики». Суть его программы состоит в допущении двойных смыслов или интенционалов у целого ряда терминов. Первичный интенционал позволяет указывать на референт в актуальном мире, вторичный — во всех возможных мирах. Скажем, первичный интенционал «воды» — «прозрачная жидкость, заполняющая океаны, реки и моря», вторичный — H_2O . Поскольку в актуальном мире прозрачная жидкость, заполняющая океаны, реки и моря — это H_2O , то свойство «быть сочетанием двух атомов водорода и одного атома кислорода» входит в понятие воды и распространяется на все возможные миры. Но если бы в актуальном мире прозрачная жидкость, заполняющая океаны, реки и моря, имела другой химический состав, все было бы иначе.

Соответственно, можно говорить о двух видах представимости: сообразно первичному интенционалу и сообразно вторичному. Так, в одном смысле воду можно представить как нечто отличное от H_2O (если внешние свойства воды могут дублироваться веществом с другой химической формулой), в другом — нет. Так же и с зомби. Зомби, по Чалмерсу, оказываются представимы сообразно первичным интенционалам качественных состояний и физических процессов, но они могут оказаться непредставимы во втором смысле в случае расхождения вторичных интенционалов физического в актуальном и воображаемом мире. Однако в чем может состоять это расхождение? Воображаемый зомби-мир, по определению, обладает той же внешней физической структурой, что и актуальный мир, и если между ними вообще могут быть отличия, они должны касаться внутренних свойств физического. Причем мыслить эти свойства различными с необходимостью можно лишь при предположении, что они совпадают с качественными или протоквалятивными состояниями. В этом случае, представляя зомби-мир, мы действительно мыслим мир, физически отличный от нашего (поскольку он мыслится с иными, некалятивными внутренними физическими свойствами), а значит, не выполняем условия мысленного эксперимента с зомби, т. е. не мыслим мир, физически тождественный нашему, но без квалиа. И тогда зомби действительно оказываются непредставимы. Но это же означает истинность неорасселовского монизма.

Таким образом, Чалмерс утверждает, что его выводы можно интерпретировать следующим образом: либо зомби представимы во всех смыслах, и тогда материализм ложен и, соот-

ветственно, истинен тот или иной дуализм, либо истинен нео-росселовский монизм. В принципе этот монизм можно истолковать в материалистическом ключе, так как качественные или «феноменальные» состояния оказываются внутренними свойствами материи. С другой стороны, сам Чалмерс подчеркивает, что неорасселовский монизм лишь номинально соответствует критериям материализма. Что же касается духа этой философии, то он скорее имма-териалистический, так как основой внешних физических свойств оказываются ментальные сущности.

В общем, может показаться, что Чалмерс находит приемлемое для себя решение: материализм в любом случае ложен, и истинен либо дуализм свойств эпифеноменалистского толка, либо имматериалистический монизм. Но здесь возникает новая сложность. Дело в том, что если мы вместе с Чалмерсом (дабы избежать противоречий внутри его системы) допускаем возможность помыслить внешне тождественный нашему мир, обладающий, однако, иными, а именно нефеноменальными, внутренними свойствами, то это закрывает перспективу решения «трудной проблемы». Ведь неорасселовский монизм, подразумевающий возможность интерпретации квалиа как основы внешних физических свойств, может выступать как вариант ее решения только когда эти качественные свойства являются безальтернативными, если другие внутренние свойства попросту невозможны. Если же это не так, то «трудная проблема» воспроизводится на новом уровне: почему функционирование нашего мозга связано именно с протофеноменальным базисом? И неясно, как решать ее.

Чтобы сохранить хотя бы перспективу решения этой проблемы, надо поэтому выбрать другой вариант: признать, что внутренними могут быть только протофеноменальные свойства, но при этом утверждать, что подобные свойства (1) качественно отличны от качественных состояний и (2) неизвестны нам. При таком толковании видимость представимости зомби оказывается связанной с неполнотой нашего знания о внутренних свойствах материи. Именно к этому варианту, который, в отличие от теории, не признающей качественного различия качественных и протофеноменальных свойств, все же может рассматриваться как разновидность теории тождества или «нейтрального монизма» (так как протофеноменальные свойства выступают в нем в качестве особой реальности, сводящей воедино физическое и ментальное), судя по всему, и склонялся Чалмерс на рубеже XX и XXI вв. Любопытно в этой связи, что в статье «Сознание и его место в природе» он прямо говорит, что такой подход сближает его взгляды с мистерианской позицией Макгинна, разобранный в одной из предыдущих глав. Само по

себе это еще ничего не значит, хотя Макгинн, напомним, считает «трудную проблему сознания» нерешаемой, а выгоды, которые можно получить от субстантивации ментального, связаны, в частности, с возможностями решения этой проблемы. В самом деле, если протофеноменальные данности качественно отличны от состояний человеческого сознания, то возникает резонный вопрос: почему и как эти протофеноменальные данности или качества порождают качественный опыт? Решение «трудной проблемы» все равно остается чем-то совершенно неопределенным.

Чалмерс и сам крайне озабочен этим скачком от протофеноменальных качеств к феноменальным, и он объявляет проблему объяснения трансформации одних в другие главной трудностью неорасселовской модели, которой он так симпатизирует. Он, правда, зачастую смещает акценты, утверждая, что основным препятствием здесь, собственно, является понимание того, как хаотичный протоопыт превращается в унифицированное человеческое сознание. Если бы все сводилось только к этому, трудности были бы устранимы. С одной стороны, мы знаем, что сознание далеко не так унифицировано, как может показаться: скорее уж оно модульно, с другой — относительная унифицированность, которая в нем все же имеется, вполне объяснима функциональным единством организма. Другое дело, если речь идет о скачке от неизвестных протофеноменальных качеств к качественным состояниям, — здесь трудности его объяснения действительно могут оказаться непреодолимыми.

Между тем можно попробовать показать, что если вообще можно отстаивать теорию базовых протофеноменальных свойств, то это можно делать только при условии того, что они непознаваемы. Вариант, когда они признаются познаваемыми и родственными качественным состояниям, просто несостоятелен. Ведь если допустить, что физические свойства несамодостаточны в силу своей релятивности и бессубстанциальности, то нетрудно убедиться, что это должно быть справедливо и относительно качественных состояний. Физические свойства известны нам из опыта, а все опытные данности могут быть интерпретированы как особого рода квалиа. А это означает, что не только физические, но и качественные свойства и состояния нуждаются в какой-то основе. Ей могут быть неизвестные протофеноменальные качества, но никак не сами качественные состояния.

Можно, таким образом, заключить, что движение Чалмерса в сторону мистерианства Макгинна — закономерный процесс. Проблема, однако, в том, что мистерианская позиция, как мы видели, сама не является логически безупречной. Рас-

суждения о непознаваемых данностях вполне могут оказаться лежащими за гранью осмысленного.

Так или иначе, теперь мы уже не будем удивляться осторожному отношению Чалмерса к перспективам обсуждавшейся выше онтологической модели, как и тому, что он в какие-то минуты признает, что все решения плохи, но если выбирать из них самое вероятное, то выбор может пасть на эпифеноменализм, который, казалось бы, преодолевается неорасселовской моделью, признающей каузальную релевантность внутренних состояний на том основании, что без них, в частности, была бы невозможна физическая причинность. Внутренние проблемы этой модели нейтрализуют ее достоинства.

Понимание всех этих обстоятельств позволяет также разгадать композиционную загадку «Сознающего ума», заключающуюся в том, что потенциальное решение «трудной проблемы сознания», стоившее Чалмерсу такого труда, вообще не подается в контексте этой проблемы! Неявным образом этот контекст, конечно, присутствует: Чалмерс обсуждает неорасселовскую модель в связи с возможностью преодоления эпифеноменализма, а анализ ранних рукописей Чалмерса показал нам, что преодоление эпифеноменализма и решение «трудной проблемы» по сути совпадают. Но тем не менее кажется удивительным, почему Чалмерс не акцентирует это. И вообще, почему первооткрыватель «трудной проблемы» так мало говорит о ней в своем главном трактате? Теперь мы знаем ответ, и он состоит в том, что Чалмерс не уверен в успешности своего решения. Соответственно, он не осмеливается выстроить композицию «Сознающего ума» таким образом, чтобы у читателей возникало предвкушение кульминации, когда «трудная проблема» была бы наконец разрешена.

Прежде чем рассказать о дальнейшей эволюции взглядов Чалмерса, я воспользуюсь подвернувшимся поводом и чуть подробнее скажу о композиции и задачах «Сознающего ума». Книга эта, несомненно, относится к числу самых значительных философских работ, созданных в последние десятилетия. Как и «Объясненное сознание» Деннета, трактат Чалмерса отличается не только оригинальностью и систематичностью, но и провокативностью — в лучшем смысле слова.

Впрочем, между этими работами, несмотря на общность тематики, есть и существенные различия. «Объясненное сознание» насыщено фактическим материалом, тогда как «Сознающий ум» — преимущественно концептуальное исследование. Кроме того, Деннет выстраивает свою книгу так, чтобы создавалось ощущение, будто он бьет в одну точку. Чалмерсу же, как мы уже поняли, тесно в рамках единственной концеп-

ции: мыслей в его книге так много, что их хватает сразу на несколько теорий, и автор колеблется между ними.

Сказанное, правда, не означает, что в «Сознающем уме» невозможно зафиксировать какой-либо четкой структуры. Наоборот, композиция этой книги очень продуманна. Помимо введения, где Чалмерс озвучивает свои главные интенции, а именно говорит о серьезном отношении к сознанию, с одной стороны, и к науке с ее тезисом о каузальной замкнутости физического — с другой, и двух приложений (посвященных проблеме искусственного интеллекта и возможным связям между квантовой механикой и теориями сознания), она включает в себя три больших раздела. В первой части, под названием «Основы», Чалмерс рассуждает о «двойной жизни» ментальных понятий и предлагает развести «психологические» (ориентированные на объяснение с точки зрения третьего лица) и «феноменальные» термины или значения одних и тех же терминов. Кроме того, он обсуждает здесь понятие «супервентности» и отличает «логическую супервентность» (логическое следование одного из другого, имеющее место, к примеру, между законами химии и физики) от «натуральной», сводящейся к корреляции неких состояний. Этим различием Чалмерс подготавливает почву для утверждения, что качественные состояния хотя и коррелятивны физическим процессам, но не могут быть редуцированы к ним в том же смысле, в каком логические следствия могут быть редуцированы к их основаниям.

Эта тема получает продолжение во второй части, озаглавленной «Нередуцируемость сознания». Чалмерс приводит аргументы, показывающие, что качественные ментальные состояния не находятся в отношении логической супервентности к мозгу или реализованным в нем функциональным схемам. Из всего этого он делает онтологические выводы о самостоятельном бытийном статусе подобных качественных состояний. Если бы некий Бог создавал наш мир, повторяет Чалмерс фразу С. Крипке, то после того, как он сотворил материю, ему пришлось бы отдельно создавать еще и сознание.

Главный из упомянутых аргументов, *de facto*, — это уже известный нам аргумент зомби⁴⁴. Я могу представить существо, тождественное мне в физическом смысле и неотличимое от меня в поведенческом плане, но при этом лишенное сознания. Если это так, то сознание как совокупность качественных состояний, утверждает Чалмерс, должно быть чем-то дополнительным по отношению к физическим и функциональным (бихевиориальным) аспектам моего Я⁴⁵.

Так Чалмерс приходит к позиции, которую он обозначает как «натуралистический дуализм». Несмотря на онтологи-

чески нагруженное название, это достаточно минималистская теория, допускающая самые разнообразные конкретизации. Ближайшей из них оказывается эпифеноменилистская трактовка натуралистического дуализма. Однако эпифеноменилизм, как мы знаем, не является беспроблемной философской теорией, и в том же втором разделе Чалмерс предпринимает немало усилий, чтобы показать, что эти проблемы не означают противоречивости эпифеноменилистских установок.

Третий раздел, «На пути к теории сознания», содержит обсуждение принципов когерентности и организационной инвариантности, которые уже были подробно проанализированы выше. Тут же Чалмерс рассматривает неорасселовскую трактовку ментального и вписывает ее в двуаспектную теорию информации, восходящую к его ранним рукописям, на которые он, кстати, здесь прямо ссылается.

Если взять «Сознающий ум» общим планом, то может показаться, что композиция этой работы все же выстроена так, чтобы создавать впечатление продвижения к решению «трудной проблемы». Уже в первых строках введения Чалмерс говорит, что, хотя «у нас есть весомое основание полагать, что сознание возникает из таких физических систем, как мозг, мы плохо понимаем, как оно возникает, да и почему вообще существует»⁴⁶. Затем, после подготовки в первой части, во второй он показывает, что трудная проблема действительно трудна. Наконец, в третьей части он ищет ее решение в двуаспектной теории информации и неорасселовской онтологии.

Но хотя эта картина и верна, она словно выполнена пунктиром и в местах обрыва линии на нее накладываются совсем другие образы. Мы уже знаем, почему это так, и я лишь подытожу сказанное ранее. Дело в том, что, если убрать из этой композиции пунктир, вся система сразу же рассыпается. Если качественное состояние действительно является фундаментальным условием физических свойств, что допускается двуаспектной онтологией Чалмерса, то, во-первых, зомби, необходимые Чалмерсу для доказательства нередуцируемости сознания и акцентуации трудности «трудной проблемы», окажутся немислимыми и невозможными, а во-вторых, удивительным образом перевернется тезис Чалмерса о порождении сознания мозгом: ведь при такой трактовке первичным будет уже сознание.

Единственный способ спасти систему от разрушения состоит в том, чтобы признать «протофеноменальность» фундаментальных свойств, совместив это допущение с тезисом об их непознаваемости, а значит, о качественном отличии этих свойств от качественных сознательных состояний. В таком случае аргумент с зомби сохраняет силу, сознание остается

нередуцируемым к физическим процессам, а «трудная проблема» не превращается в легкую. Не происходит эксцессов И с положением о порождаемости качественных состояний мозгом. Но ценой станет то, что «трудная проблема» не будет решена.

В самом деле, вопрос о том, почему мозг порождает сознание, так и останется без ответа. Мы получим ответ на вопрос, почему с мозгом связаны протофеноменальные, протоквали-тативные состояния, но загадка будет состоять в том, почему эти неведомые протофеноменальные состояния вместе с мозгом генерируют наш сознательный опыт. Конечно, можно допустить, что, если бы мы узнали получше эти протофеноменальные свойства, мы смогли бы ответить на этот вопрос. Макгинн, с которым, как мы уже отмечали, Чалмерс готов вступить в союз (хотя Макгинн пользуется другой терминологией), считает, правда, что эти свойства в принципе непознаваемы. И это тоже вариант решения трудной проблемы, пускай и сугубо отрицательный. Пикантность ситуации, однако, в том, что Чалмерс не разделяет радикализма Макгинна. Он не утверждает, что протофеноменальные состояния не могут быть познаны. Чалмерс полагает, скорее, что этот вопрос еще открыт. Но в таком случае он вообще оставляет «трудную проблему» без решения. Фактически он просто констатирует, что «трудная проблема» трудна, а вот можно ли ее решить — неизвестно. Точно такую же позицию, как мы помним, занимал Сёрл, она характерна и для австралийского философа Д. Столяра⁴⁷.

Ситуация выглядит несколько парадоксально, и едва ли можно предположить, что Чалмерс мог остановить свои поиски на столь неопределенной ноте, к тому же затушеванной многочисленными оговорками, делающими «Сознающий ум» лабиринтом для любого интерпретатора. Логично поэтому еще раз взглянуть на его более поздние работы, и прежде всего на статью «Сознание и его место в природе». Внимание к ней оправдывает себя хотя бы потому, что Чалмерс критикует здесь некоторые положения «Сознающего ума».

Об этой критике я скажу чуть позже, пока же отмечу, что заглавие статьи — *Consciousness and Its Place in Nature* — специально подобрано Чалмерсом так, чтобы оно перекликалось с названием книги Ч.Д. Броуда *The Mind and Its Place in Nature* (1925). Чалмерс вообще пытается восстановить интерес к этому кембриджскому мыслителю, основательно подзабытому в наши дни. Широкая философская публика знает о Броуде лишь то, что он был недоброжелателем Витгенштейна, большим педантом (говорят, что для закрепления знаний у студентов он произносил каждую фразу своих лекций два-

ды, а шутки — трижды) и «эмерджентистом». Я не буду вдаваться в детали его труда о сознании и хочу лишь отметить, что симпатия Чалмерса к Броуду во многом связана с тем, что Чалмерс, как и Броуд, не чужд классификаторского азарта. И хотя Чалмерс не дотягивает до 17 вариантов решения вопроса о соотношении сознания и тела, о которых говорит Броуд, в статье «Сознание и его место в природе» он тоже предлагает детальную панораму «трудных ответов».

Чалмерс выделяет шесть принципиальных подходов. Три из них — материалистические, остальные — нередуктивистские. Материалисты могут либо элиминировать сознание, либо признавать его реальность как предмета познания, но при этом настаивать на его неразрывной связи или онтологическом тождестве с чем-то физическим, либо, наконец, утверждать, что это тождество когда-нибудь будет установлено. Первый вариант, по мнению Чалмерса, противоречит фактам и поэтому неприемлем (с этим нельзя не согласиться), а третий нестабилен и при его конкретизации коллапсирует в одну из альтернативных теорий. Что же касается второго варианта, отстаиваемого теоретиками тождества, то он является, как считает Чалмерс, более реальной опцией и заслуживает подробного анализа.

Я, однако, не разделяю эту позицию и уже высказывался на этот счет. Теория тождества не допускает верификации и может быть сразу отвергнута. Но Чалмерс считает, что для ее отклонения требуется основательная аргументативная подготовка. Впрочем, его особое внимание привлекает, скорее, даже не теория тождества, а та разновидность материалистического решения проблемы сознания, при которой признается «метафизически необходимая» связь между качественными и физическими состояниями. Сторонники этого учения апеллируют к некоторым идеям С. Крипке, в частности к его тезису о существовании так называемой апостериорной необходимости.

Самый известный пример крипкеанской апостериорной необходимости: «Вода есть H_2O ». То, что вода есть H_2O , мы узнаем из опыта, т. е. апостериори, но при этом трудно отрицать, что вода должна иметь эту химическую структуру во всех возможных мирах, истинность же некоего тезиса во всех возможных мирах и означает его необходимость. И возникает вопрос: а нельзя ли применить этот подход и к проблеме сознания? В самом деле, не может ли быть так, что, хотя мы не в состоянии априори показать, что сознание необходимо связано с физическими процессами, эта связь тем не менее имеет характер апостериорной необходимости, из чего, в частности, следует непредставимость зомби и отрицание тезиса о нереду-

цируемости сознания, являющегося главным доводом против материализма.

Эти аргументы могут показаться опасными, и Чалмерс использует уже упоминавшийся аппарат «двумерной семантики», чтобы отвести их и показать, что в случае сознания метафизическая необходимость его связи с мозгом может рассматриваться как реальная опция только при неорасселовской интерпретации ментального. Но не исключено, что теорию Крипке и ее возможные следствия для трактовок сознания можно нейтрализовать и другим способом. Вполне возможно, что само понятие апостериорной метафизической необходимости — голый король. Мне трудно понять, к примеру, почему Крипке считает, что высказывание «вода есть H_2O » апостериорно. Его вполне можно истолковать как аналитическое, а значит, априорное. Конечно, для того, чтобы включить предикат « H_2O » в понятие воды, потребовались научные открытия, и были времена, когда данное утверждение не было аналитическим. Но вообще-то все аналитические утверждения - в классическом кантовском смысле - генетически связаны с опытом⁴⁸. Однако это нерелевантно для их статуса. Понимание этого ставит под вопрос теорию Крипке⁴⁹.

Так или иначе, но, отвергнув материалистические решения, Чалмерс констатирует, что у нас остается три нередуцируемых варианта трактовки статуса сознания, и он тщательно рассматривает их особенности. Первый — это интеракционистский дуализм, второй - эпифеноменализм, третий - неорасселовский монизм, при котором (прото)феноменальные состояния оказываются основой физического.

И вот здесь обнаруживается явное отличие рассматриваемой статьи от «Сознающего ума». Чалмерс отказывается от радикальной критики интеракционизма и, соответственно, от тезиса о каузальной замкнутости физического как единственного базиса перспективных теорий сознания. Констатируя, что большинство философов отвергают интеракционизм на том основании, что он якобы не стыкуется с физикой, он добавляет: «Я и сам был повинен в этом, отставляя интеракционизм в моей книге 1996 г. отчасти в связи с совместимостью с физикой. Я и сейчас не особо склонен одобрять интеракционизм, но теперь я считаю, что аргумент по поводу физики не слишком серьезен»⁵⁰. И в самом деле, согласно некоторым интерпретациям квантовой механики, так называемый коллапс волновой функции, всякий раз происходящий при фиксации эмпирических данностей, невозможен без участия наблюдателя, причем сознательного наблюдателя, так как измерительные приборы как таковые не могут вызывать этот коллапс, будучи сами подвержены суперпозиции, т. е. словно

бы находясь в разных состояниях одновременно. Если это действительно так, то сознание играет определенную каузальную роль, актуализируя один из множества вариантов развития событий.

Разумеется, все эти вещи были известны Чалмерсу и ранее. В «Сознающем уме» он отвергает интерпретации, отводящие сознанию реальную каузальную роль в коллапсе волновой функции, в частности или даже прежде всего потому, что они «несовместимы со взглядом о повсеместности сознания, который я защищал. Если сознание связано даже с очень простыми системами, то, согласно этой интерпретации, коллапс будет происходить на самом базовом уровне и очень часто. Это противоречит физическим данным, зачастую подразумевающим неколлапсированное существование низкоуровневых суперпозиций в течение продолжительного времени»⁵¹.

Неудивительно, что в «Сознающем уме» Чалмерс отдавал предпочтение (хотя и с многочисленными оговорками) интерпретации Эверетта⁵², в соответствии с которой никакого коллапса вообще не происходит, а происходит то, что мир эволюционирует сразу во множестве направлений, в одном из которых случается быть нам. Эта интерпретация более или менее сочетается с принципом каузальной замкнутости физического, но Чалмерс также утверждал, что она гармонирует с выведенными им психофизическими принципами.

О предпочтительности интерпретации Эверетта Чалмерс говорил и в статье «Продвигаясь к проблеме сознания», опубликованной через год после выхода «Сознающего ума». Чалмерс шутит здесь, что четыре дня в неделю он является сторонником Эверетта, но два дня он все же отводит на интеракционистские интерпретации квантовой механики, а воскресенья оставляет на интерпретацию, связанную с введением скрытых параметров⁵³.

Как видим, изменение отношения к интеракционизму наметилось у Чалмерса уже в 1997 г., и эта тенденция еще сильнее проявилась в статье «Материализм и метафизика модальности», написанной в 1998 г. В статье 1997 г. он, правда, отмечал, что даже признание каузальной действительности сознания не избавляет от опасности эпифеноменализма, так как его каузальная роль все равно могла бы быть вынесена за скобки, да и сама сознательная природа этой каузальности выглядит чем-то случайным: вполне можно помыслить ситуацию, когда эта каузальная роль выполнялась бы чисто физическими агентами. Это возражение требует более детального рассмотрения, чем может показаться, и я еще вернусь к данной теме в следующей главе. Сейчас же отмечу, что, даже если признать мыслимость сценариев, описываемых Чалмерсом,

нас в любом случае интересует не возможный, а реальный статус сознания⁵⁴. Поэтому, даже если каузальная роль сознания акцидентальна, знание о том, что оно играет эту роль, все равно теоретически значимо, и поэтому приведенное выше возражение Чалмерса не имеет большого веса.

Чалмерс, вероятно, и сам почувствовал неоднозначность этого возражения, и в статье «Сознание и его место в природе» он даже не упоминает о нем. Уходит в тень в этой статье 2002 г. и интерпретация Эверетта. С той же энергией, с какой он раньше говорил о ней, он рассуждает об интерпретациях, подразумевающих участие сознания в коллапсе волновой функции. И хотя он не объявляет себя интеракционистом, а говорит, скорее, о равновероятности интеракционизма, эпифеноменализма и неорасселовского монизма, тем не менее именно интеракционистская позиция поглощает его внимание.

Что же случилось? Почему Чалмерс сместил акценты в сторону интеракционизма? Ведь, согласно приведенным чуть выше утверждениям самого Чалмерса, он противоречит теории повсеместности психического, являющейся естественным следствием неорасселовского монизма, к которому он тяготел в «Сознающем уме». Ответ, вероятно, состоит в том, что Чалмерс осознал, что он не имеет права апеллировать к этой монистической теории как к независимому доводу против интеракционистских интерпретаций квантовой механики — потому что сама эта теория, как мы знаем, была разработана им при предположении каузальной замкнутости физического. Если он, таким образом, не хотел обрекать себя на круг в аргументации, он должен был заново и отвлекаясь от всяких внешних допущений сопоставить интеракционизм и неорасселовский монизм. И сопоставление это выявило то, что интеракционизм свободен от главной, с его точки зрения, проблемы неорасселовского монизма, состоящей в трудности объяснения перехода от протофеноменальных свойств, фундирующих все физические процессы, к качественным состояниям.

Единственный ясный способ, по Чалмерсу, уйти от этой проблемы — объявить, что феноменальных состояний на микроуровне (протофеноменальных состояний) просто не существует и что феноменальные состояния являются внутренними свойствами высокоуровневых материальных систем. Но тогда, чтобы избежать эпифеноменализма, на преодоление которого была нацелена его неорасселовская модель, надо допустить, что физическая реальность не является замкнутой на микроуровне. В таком случае феноменальные, качественные состояния смогут оказывать влияние на физические процессы. Но эта теория, как подчеркивает сам Чалмерс, уже ни-

чем не отличается от интеракционизма. Таким образом, мы можем констатировать объективную тенденцию эволюции идей Чалмерса к интеракционистскому взгляду на каузальную роль сознания: неорасселовское учение о квалитативном или феноменальном как внутреннем аспекте физического объективно лучше эпифеноменалистской онтологии, но само оно выглядит наиболее защищенным от критики при отрицании протофеноменальных состояний, которое, однако, является антитезой эпифеноменализму лишь при условии признания каузальной действительности квалитативных состояний, что, в свою очередь, приводит к трансформации данной теории в интеракционизм. Не будем забывать, правда, что это лишь тенденция и официально Чалмерс отдает должное всем нередукционистским концепциям.

Эта онтологическая политкорректность Чалмерса, которую, как мне кажется, читатель или читательница уже прочувствовали, не могла рано или поздно не привести его к чему-то большему. Ведь если мы постоянно говорим, что может быть истинна и эта, и та позиция и что мы не знаем, как выбрать между ними, или каждый день выбираем из этого онтологического меню то, чего нам больше хочется сегодня, то мы невольно должны будем задумываться о том, с чем это связано и какая метапозиция допускает такие действия. В 2007 г. Чалмерс определенно ответил на эти вопросы, заявив, что он является сторонником онтологического антиреализма, установки, согласно которой свойства мира не есть что-то абсолютное: они зависят от эпистемологических контекстов⁵⁵. Так осторожность в решении одних проблем оборачивается экстремизмом в других, и в этом смысле совсем не случайно, что, пройдя один из мировоззренческих онлайн-тестов, Чалмерс обнаружил, что он «постмодернист»: постмодернизм и эпистемологический релятивизм — почти одно и то же.

Разумеется, эпистемологический релятивистский экстремизм и его онтологический спутник, антиреализм, — это позиция, мало подходящая для преодоления «трудной проблемы». Между тем возможно, что Чалмерс отказался от попыток ее решения в тот самый момент, когда ближе всего подошел к разгадке, проявляя все большие симпатии к интеракционизму. Ведь хотя из текстов Чалмерса трудно вычитать указания на то, как интеракционистская позиция позволяет объяснить, почему функционирование мозга вообще сопровождается сознанием, в следующей главе мы увидим, что у сторонников этой доктрины имеются некие ресурсы для ответа на данный вопрос.

Примечания

¹ Из этого правила есть исключения, к примеру: Pauen M. Was ist der Mensch? Die Entdeckung der Natur des Geistes. Munchen, 2007. Но оно подтверждает правило, так как автор, не упоминая Чалмерса, пользуется его идеями (см., напр.: S. 130-132). Вообще же философы благодарят «Дэйва за то, что он Дэйв», и говорят, что, если бы его не было, его надо было бы придумать.

² В юности Чалмерс был победителем Всеавстралийской олимпиады по математике, выиграл еще несколько национальных конкурсов и занял призовое место на международной олимпиаде.

³ В своей книге «Я — это странная петля» Хофштадтер, этот самый романтичный из американских философов (перевод им «Евгения Онегина» явно не был случайностью), сокрушенно вопрошает, как же под его крылом мог вырасти философ, столь чуждый ему по своим взглядам и умонастроениям (сам Хофштадтер высказывает те же идеи, что и его друг Деннет, но только еще более метафорично). См.: Hofstadter D. I am a Strange Loop. N. Y., 2007. P. 375.

⁴ Официально (см. сайт туссанского Центра исследований сознания, раздел о конференции 2006 г.) первые строки его таковы: «I act like you act, I do what you do, but I don't know, what it's like to be you. What consciousness is, I ain't got a clue. I got the Zombie Blues!...».

⁵ Здесь и далее см.: <http://consc.net/notes/first-third.html>

⁶ См.: <http://consc.net/papers/c-and-c.html>

⁷ Ibid. Сходную теорию разрабатывает российский философ Д.И. Дубровский. См.: Дубровский Д. И. Сознание, мозг, искусственный интеллект. М., 2007. С. 114-163. В отличие от Чалмерса, однако, Дубровский утверждает, что квалиа, субъективные состояния, обладают реальным каузальным воздействием на организм — и это воздействие является одним из примеров так называемой «информационной причинности». Главная проблема теории Дубровского, по-видимому, в том, что он не вполне проясняет, есть ли информационная причинность просто высокоуровневое проявление базовой физической причинности, или же она представляет собой нечто самостоятельное. Иными словами, являются ли каузальные возможности квалиа чем-то превосходящим каузальные возможности нейронных «кодов», в которых они реализованы, — и если да, то как можно согласовать это с принципом каузальной замкнутости физического. Между тем создается впечатление, что он склоняется к отрицательному ответу. Но такой ответ крайне затрудняет объяснение того, зачем вообще нужны квалиа. Попытки подобного объяснения неизбежно возвращаются по кругу, поскольку всегда можно сказать, что все те функции, с которыми предположительно связаны квалиа, могли бы быть реализованы и без них.

⁸ Elitzur A. Consciousness and the incompleteness of the physical explanation of behavior // Journal of Mind and Behavior 10 (1989). P. 1-20. Хотя Элитцур и отсылал к сходным рассуждениям Р. Пенроуза в статье «Квантовая физика и осознанное мышление» (1987), но, поскольку они были очень уж краткими (Пенроуз утверждал,

что «если бы сознание не оказывало действенного влияния на поведение, сознающие существа никогда не выказывали бы своего удивления по поводу сознательного состояния и вели бы себя как бессознательные механизмы, которых "не беспокоит" подобная чепуха» — Penrose R. *Quantum physics and conscious thought*//*Quantum Implications: Essays in Honour of David Bohm*, ed. by B. J. Hiley, F. David Peat. L., 1991 (1987). P. 116), впоследствии он счел возможным представить дело так, будто это доказательство действенности сознания является его собственным изобретением — см.: Elitzur A. *Consciousness makes a difference: Reluctant dualist's confession*. TS 2005. В действительности используемый Элитцуром аргумент (восходящий, помимо прочих менее известных источников — см.: Koks-vik O. In *Defence of Interactionism*. Master thesis. TS 2006. P. 26 -к работам Б. Мэдлина и С. Шумейкера (если, конечно, не принимать во внимание его аналог, обсуждавшийся еще в школе Вольфа, -см., напр.: *Sammlung und Auszuge der sammtlicheu Streitschriften wegen der Wolffischen Philosophie, zur Erlauterung der bestritten Leibnitzischen und Wolffischen Lehrsatze...* verfertigt von Carl Guntor Ludovici. Th. 1. Lpz., 1737. S. 195), в частности — как указывал Плейс - к статье Мэдлина «Райл и механическая гипотеза» (1967), а также к статье Шумейкера «Функционализм и квалиа» (1975) — см.: Shoemaker S. *Functionalism and qualia* // *The Nature of Mind*. P. 397-398) был более или менее проработан в 80-е годы. Так, он встречается в статье Д. Льюиса «Чему учит опыт», впервые опубликованной в 1988 г., — см.: Lewis D. *What experience teaches* // *There's Something About Mary*. P. 94-95. Его использует и Р. Суин-берн в «Эволюции души», первое издание которой появилось в 1986 г., — см.: Swinburne R. *Evolution of the Soul*. P. 40(Суинберн был знаком с текстами Шумейкера). Льюис, как и Плейс, в том же году рассуждавший на ту же тему — см.: Place U. T. *Identifying the Mind*. P. 79, — реагирует на него бегством в материализм. Суин-берн, как и Элитцур, предпочел ему интеракционистский дуализм (хотя в более поздних статьях Элитцур все больше акцентирует и другой вариант: возможность того, что наши высказывания о собственном сознании могут быть результатом когнитивных ошибок, — причем эту гипотезу, утверждает он, можно было бы проверить экспериментально). Чалмерс, как мы увидим, избрал третий путь.

⁹ Jackson F. *Epiphenomenal qualia*. P. 48. Джексон считал, что мы не в состоянии ответить на эти вопросы.

¹⁰ На титульном листе диссертации указана дата защиты — 29 ноября. Я благодарен Д. Чалмерсу за предоставленную возможность ознакомиться с текстом его диссертационного исследования.

¹¹ Хотя имплицитно он содержится уже в некоторых рассуждениях Декарта. К примеру, в письме Г. Мору от 5 февраля 1649 г. Декарт дает понять, что считает более удивительным наличие сознания у человека, чем его отсутствие у животных, — см.: Декарт Р. *Сочинения* в 2 т. Т. 2. М., 1994. С. 575. Такие высказывания вполне можно переформулировать в вопросительной форме — и мы получим «трудную проблему». Более того, у Декарта можно найти и вариант ее решения. Он считал, что автоматы, в том числе животные, не способны продуцировать речь, проходить, как мы бы сказали,

тест Тьюринга. Впрочем, Декарт, что называется, недостаточно раскрыл эту тему. Вопрос «Почему функционирование мозга сопровождается субъективным опытом?» также близок вопросу о функции сознания. Подобный вопрос мы встречаем, к примеру, в статье У. Джемса «Являемся ли мы автоматами?» — см.: James W. Are we automata? // *Mind* 4 (1879). P. 1-22.

¹² 1994 г. выражение «трудная часть проблемы сознание—тело» в сходном с чалмерсовским смысле использовал также Г. Стросон — см.: Strawson G. *Mental Reality*. Cambridge, 1994. P. 93. Ср.: Block N. The harder problem of consciousness // *The Journal of Philosophy* 99: 8(2002). P. 391.

¹³ Вместо этого, следуя Р. Джекендоффу, он пользуется выражением «проблема сознание-сознание» (mind-mind problem), имея в виду проблему того, почему сознание как функциональная схема (или сознание в психологическом смысле) порождает феноменальное сознание. Это место осталось и в «Сознающем уме».

¹⁴ Письмо от 27 декабря 2007 г.

¹⁵ Chalmers D.J. Facing up to the problem of consciousness. P. 9.

¹⁶ Ibid. P. U.

¹⁷ Ibid. P. 13.

¹⁸ Ibid. P. 20.

¹⁹ Написана в 1996 г.

²⁰ Chalmers D. J. Moving forward on the problem of consciousness // *Explaining Consciousness*. P. 400.

²¹ Об этом пишет и сам Чалмерс. См.: Chalmers D. J. On the search for the neural correlate of consciousness // *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates*, ed. by S. R. Hameroff, A. W. Kaszniak, A. C. Scott. Cambridge MA, 1998. P. 219-230.

²² Chalmers D. J. *The Conscious Mind*. P. 276.

²³ Chalmers D. J. Facing up to the problem of consciousness. P. 22.

²⁴ Здесь «Сознающий ум» следует диссертации 1993 г.

²⁵ Chalmers D. J. *The Conscious Mind*. P. 244.

²⁶ Ibid.

²⁷ Чалмерс прямо говорит об этом в диссертации — см.: Chalmers D. J. *Toward a Theory of Consciousness*. Ph. D. Thesis. TS 1993. P. 202. В «Сознающем уме» это место сокращено — P. 249.

²⁸ Опубликована в 1995 г.

²⁹ См.: Ibid. P. 247-274.

³⁰ Чалмерс отмечает, что к этому аргументу его подтолкнула история «Где Я?», рассказанная Деннетом в книге «Глаз разума» — см.: Chalmers D. J. *The Conscious Mind*. P. 388.

³¹ Ср.: Koch C *The Quest for Consciousness*. P. 99.

³² См.: *Philosophy and Phenomenological Research* 59 (1999). P. 439-44.

³³ Ibid. P. 473-493.

³⁴ Ibid. P. 493.

³⁵ Тем более что субъективная яркость вообще является, по сути, относительным феноменом.

³⁶ Chalmers D. J. Consciousness and its place in nature // *Philosophy of Mind: Classical and Contemporary Readings*, ed. by D. J. Chalmers. N. Y., 2002. P. 264.

³⁷ В статье «Навстречу проблеме сознания» (1995) он вообще обходит этот вопрос стороной.

³⁸ Chalmers D. J. *The Conscious Mind*. P. 160.

³⁹ *Ibid.* P. 277-283.

⁴⁰ Bateson G. *Steps to an Ecology of Mind*. Chicago, 2000. P. 459.

⁴¹ Chalmers D. J. *The Conscious Mind*. P. 305.

⁴² Издавшего, однако, свой трактат — «Место сознания» (2004) — гораздо позже «Сознающего ума» Чалмерса. В книге Розенберга выражена та же позиция, что рассматривается далее, но стиль этого автора отличается удивительной непрозрачностью. И поскольку я считаю его онтологическую модель малоперспективной, я обойду стороной эту любопытную монографию.

⁴³ См.: Russell B. *The Analysis of Matter*. New ed. N. Y., 1954. P. 384, 400, 402. Эта онтологическая конструкция, в которую эволюционировал нейтральный монизм Рассела, заставляет вспомнить уже не Джемса, который повлиял на его формирование, а Канта и Шопенгауэра.

⁴⁴ Часто его называют «аргументом представимости». Близкий ему и рассмотренный нами ранее аргумент Ф. Джексона («Нейро-ученая Мэри») относится к группе «аргументов знания».

⁴⁵ Этот аргумент (без употребления термина «зомби») в четком виде был сформулирован уже Т. Нагелем в 1970 г. См.: Nagel T. *Other Minds: Critical Essays*. 1969-1994. N. Y., 1995. P. 79.

⁴⁶ Chalmers D. J. *The Conscious Mind*. P. XI.

⁴⁷ См.: Stoljar D. *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness*. N.Y., 2006. В этой — композиционно отполированной и тщательно выписанной — книге Столяр утверждает, что противоречия, возникающие при попытке мыслить отношения сознания и мозга, являются следствием нашего незнания неких «вне опытных истин», т. е. истин, не имеющих отношения к фактам сознания. В отличие от Макгинна Столяр не утверждает, что мы когнитивно замкнуты по отношению к этим истинам. И он полагает, что его позиция позволяет разрешить «логическую проблему опыта», возникающую вследствие некогерентности наших естественных убеждений, касающихся (1) существования «опытных истин», (2) их зависимости от вне опытных истин (физикализм) и (3) их независимости от подобных истин (очевидной из зомби-аргументов). Все это замечательно, но я совершенно не понимаю, почему Столяр считает, что данная теория показывает, что философия сознания как концептуальная программа, позволяющая разрешить «метафизические проблемы сознания и, в частности, опыта», «не имеет перспектив» (р. 233). В действительности его минималистская позиция сводится к утверждению, что, хотя сознание наверняка является продуктом мозга, нам еще неизвестны те свойства мозга, которые с необходимостью порождают сознание. Но из этого не следует, что с помощью концептуального анализа невозможно показать, о какого рода — или даже о каких — свойствах идет речь. Иными словами, Столяр подтверждает, что «трудная проблема сознания» действительно трудна, но подобное подтверждение не может заблокировать поиск ее концептуального решения.

⁴⁸ Разумеется, не надо забывать, что они не совпадают с пустыми тавтологиями.

⁴⁹ Ключевой вопрос в связи с этой теорией — вопрос о «жестких десигнаторах», к числу которых, по Крипке, относится и термин «вода». На мой взгляд, Крипке не обратил внимания на то, что все высказывания, включающие жесткие десигнаторы, т. е. термины, связанные со своими референтами во всех возможных мирах, являются высказываниями о самих терминах, т. е. о правилах их корректного употребления, а не об обозначаемых ими предметах. Между тем понятие метафизической необходимости, имплицитруемой жесткими десигнаторами, имеет предметный смысл. Таким образом, это понятие совершенно неправомерно и должно быть отброшено. Разумеется, я отдаю себе отчет, что порой фраз нельзя разделить со столь влиятельной теорией. И все же я не буду вдаваться в детали — это уведет нас от нашей основной темы. Скажу лишь, что современная философия языка, на мой взгляд, нуждается в глубоком реформировании, которое надо начинать с чистого листа — экстенсивное развитие здесь затруднено.

⁵⁰ Chalmers D. J. *Consciousness and its place in nature*. P. 270.

⁵¹ Chalmers D. J. *The Conscious Mind*. P. 340.

⁵² В диссертации 1993 г. он поддерживал ее гораздо более решительно.

⁵³ Chalmers D. J. *Moving forward on the problem of consciousness*. P. 403.

⁵⁴ Ср.: Polger T. *Rethinking the evolution of consciousness* // *The Blackwell Companion to Consciousness*. P. 78. Полгер справедливо утверждает, что из тезиса о «несущественности» сознания для реализации той или иной функции (т. е. из тезиса о том, что эта функция могла бы быть реализована и при каких-то других условиях) еще нельзя заключать к его эпифеноменальности. См. также: Flanagan O. *Conscious inessentialism and the epiphenomenalism suspicion* // *The Nature of Consciousness*. P. 357-373.

⁵⁵ См.: Chalmers D. J. *Ontological anti-realism*. TS 2007.

Глава 5 НАЗАД В БУДУЩЕЕ

Прежде чем двигаться дальше, нам надо осмотреться и оценить ситуацию, в которой мы оказались. Внимательные читатели могли заметить, что она несколько отличается от той, перспектива которой вырисовывалась в конце первой главы. Там шла речь о том, что в последующих главах будет представлена почти что гегелевская триада решений «трудной проблемы сознания» и что мы увидим диалектическое продвижение от прямолинейной «тетической» установки Сёрла, раздувающего «квалитативные» интуиции здравого смысла в пику когнитивной науке, которая изучает сознание с позиции третьего лица как свойство определенных физических систем, к абсолютному квалитативному нигилизму когнитивиста Деннета, отрицающего само существование сознания, и, наконец, завершим нашу экспозицию синтетическими схемами Чалмерса, в которых ментальное все же будет приведено к общему знаменателю с физическим. Действительность оказалась менее драматичной по гегелевским стандартам, но существенно более драматичной в другом плане.

И драматизм этот заключается в том, что пока мы просто не увидели реальных попыток позитивно решить «трудную проблему». Конечно, мы прошли по галерее онтологических схем соотношения квалитативных состояний и физических процессов в мозге, но ни одна из них не позволяла четко ответить на вопрос, почему эти физические процессы вообще сопровождаются сознанием, а не происходят, по удачному выражению Чалмерса, «в темноте». Между тем именно этот вопрос составляет ядро «трудной проблемы». Поскольку даже сам изобретатель «трудной проблемы» воздерживается от ее решения, то нельзя не задуматься — а стоило ли вообще ее поднимать? Или это стоило сделать для того, чтобы кто-то задался целью показать, что она в принципе не может быть решена? Такой ограничительный проект, напомним, пытался реализовать К. Макгинн. Выходит, что он единственный, кто всерьез хотел решить «трудную проблему», с тем, правда, уточнением, что он не использовал термина «трудная проблема сознания» и что его решение чисто негативно¹.

Не будем спешить с ответами. И не будем отбрасывать рассмотренные ранее позитивные теории — я анализировал их совершенно не для того, чтобы в конце концов торжественно провозгласить, что они ничего не дают. Наоборот, я считаю,

что именно эти теории, как ничто иное, могут подтолкнуть к решению «трудной проблемы». Эту роль они могут сыграть потому, что «трудная проблема» как таковая связана с другими важными вопросами, которые обсуждаются и решаются авторами упомянутых теорий. Одним из них является вопрос об онтологическом статусе квалитативных состояний. Так что забудем на время о «трудной проблеме» в ее, так сказать, «ядерном» аспекте и займемся ее периферийными моментами — в надежде, что нам еще удастся вернуться к нашей главной теме.

Вопрос об онтологическом статусе квалитативных состояний, на котором мы должны будем сосредоточиться, может, правда, показаться очень уж размытым. Отчасти с этим можно согласиться, но в любом случае у него есть один ракурс, который допускает конкретное рассмотрение. Речь идет о каузальной роли квалиа по отношению к физическим процессам. Очевидно, что понимание этой роли напрямую увязано с характеристикой онтологического статуса ментальных состояний. Если, к примеру, квалиа каузально бесплодны, то они либо вообще не существуют, либо должны быть признаны подозрительными эпифеноменами. С другой стороны, если можно показать, что квалиа обладают реальной действенностью (рассуждения о каузальной релевантности неких протоквалитативных состояний, как я показал раньше, лишены всякого смысла), то перед нами открываются совершенно иные перспективы. Квалиа могут оказаться нередуцируемой, а возможно даже субстанциальной реальностью, вторгающейся в материальные процессы и разрушающей каузальную замкнутость физического. Если же мы держимся за принцип каузальной замкнутости, но при этом хотим признать каузальную релевантность квалитативных состояний², то нам, как считали многие, ничего не остается, как решиться на радикальное отождествление квалитативных состояний и физических процессов в мозге.

Урок предыдущих глав состоит, однако, в том, что все перечисленные варианты выглядят очень сомнительно. Бесперспективно элиминировать квалитативные состояния, как это делает Деннет, но бесперспективно и признавать их самодостаточность — слишком уж очевидна фундированность таких состояний нейронными процессами. Да и отрицание каузальной замкнутости физического — слишком серьезный шаг, чтобы делать его без необходимости. Впрочем, если не делать его, то, как кажется, надо выбирать между теорией тождества, которая, как я пытался показать ранее, в принципе неverifiedируема и в силу этого несостоятельна, и эпифеноменализмом³. Соответственно, приходится выбирать эпи-

феноменализм. Мало кто спорит, что эпифеноменализм — неуклюжая конструкция, но теперь мы можем понять, почему Чалмерс (и это тоже важный урок) писал, что, хотя все решения проблемы сознания плохи, эпифеноменалистское может оказаться лучшим из того, что имеется у нас в наличии.

В самом деле, эпифеноменализм по крайней мере не отрицает тезис о реальности квалиа. Кроме того, он не отрицает каузальную замкнутость физического, на допущении которой, как многие полагают, держится чуть ли не все естествознание. Платить за совмещение этих тезисов приходится тем, что сторонники эпифеноменализма вынуждены оспаривать одну из аксиом здравого смысла, а именно представление о том, что наши сознательные состояния оказывают влияние на поведение. Нейтрализовать эту аксиому, впрочем, несложно. Здравый смысл говорит нам, что если я, скажем, хочу поднять руку, то желание сделать это является по меньшей мере одним из компонентов причины этого моего последующего действия. Эпифеноменализм, однако, настаивает на отсутствии каузального потенциала у подобных ментальных состояний. Противоречие, впрочем, можно устранить, отметив, что реальной причиной поднятия руки может быть нейронное событие, скоррелированное с желанием сделать это. Здравый смысл попросту не учитывает существования такого события, но он не может не согласиться с возможностью подобного объяснения. Иными словами, феноменология корреляции желания поднять руку и соответствующего физического движения не меняется от того, истинен или ложен эпифеноменализм. Поскольку здравый смысл базирует свои выводы о реальности ментальной каузальности только на этой феноменологии, здравый смысл не может рассматриваться как довод против эпифеноменалистской позиции⁴.

В тех местах «Сознающего ума», где Чалмерс защищает эпифеноменализм, он, разумеется, рассматривает подобный ответ критикам эпифеноменализма с позиций здравого смысла. Ответ, заметим, вполне убедительный⁵. Проблема, однако, в том, что эпифеноменализм можно критиковать и с других позиций. Чтобы разобраться в этом, уместно напомнить историю этой любопытной концепции.

Традиционно считается, что первым последовательным эпифеноменалистом был Томас Гексли. Он, правда, не употреблял термина «эпифеноменализм», используя другое название — «автоматная теория»⁶. Но, вообще говоря, тезис о каузальной бесплодности психических состояний отстаивал еще Хр. Вольф. Мотивы Вольфа, впрочем, отличались от тех, что принимал в расчет Гексли в конце XIX века. Вольф отталкивался от лейбницевской идеи предустановленной гар-

монии между душой и телом, в свою очередь зависящей от положения, что наш мир является наилучшим из возможных миров. Согласно концепции предустановленной гармонии, психические и телесные состояния коррелятивны, но реально не влияют друг на друга. Самого Лейбница, учителя Вольфа, трудно считать эпифеноменалистом (хотя Спиноза, подтолкнувший Лейбница к теории предустановленной гармонии, был близок эпифеноменализму), так как его концепция физического не укладывается в привычные контексты проблемы сознание—тело: тела в основе своей являются совокупностями психических монад, и в рамках его системы можно, скорее, говорить о гармоническом соответствии перцепций этих простых субстанций. Вольф же лишил телесные сущности психических характеристик и поэтому оказался гораздо ближе современному эпифеноменализму.

В отличие от Вольфа, эпифеноменализм которого базировался на метафизических предпосылках — и был все же относительным, так как классический эпифеноменализм говорит не только о каузальной бесплодности ментального по отношению к физическому, но и о том, что сознание порождается мозгом, а Вольф это отрицал (поэтому точнее относить его, как и Спинозу, к параллелистам)⁷, — Гексли отсылал к естественнонаучным данным, свидетельствующим о независимости телесных движений от наличия или отсутствия сознания. Обобщив имеющийся у него экспериментальный материал относительно поведения животных и людей с повреждениями мозга, он заключил, что психические состояния не влияют на поведение, являясь чем-то побочным, чем-то таким, что можно сравнить с паровозным свистком. Подобно пару от свистка, эти состояния порождаются мозгом, но не оказывают на него какого-либо обратного влияния. Впрочем аналогия неточна. Пар от свистка вследствие гравитации все же оказывает влияние — пусть и ничтожное — на движение паровоза. Эпифеноменалисты отрицают даже ничтожное влияние⁸.

Отметим, что эпифеноменализм — в варианте, предложенном Гексли, — не выглядит хорошо обоснованной доктриной. Из того, что лягушки могут совершать сложные движения при удалении частей головного мозга — на этот факт, в частности, и ссылался Гексли, — еще не следует, что поведение организма не нуждается в сознании. Ведь роль сознания, к примеру, может состоять в корректировке или «тонкой настройке» рефлекторных поведенческих реакций. Это предположение подтверждается анализом интереснейших экспериментов, сделанных в последние десятилетия и связанных, в частности, с исследованием феномена так называемого «слепого зрения»⁹.

Слепое зрение обнаруживается у людей и животных¹⁰ с поврежденным затылочным участком коры головного мозга (V1), отвечающим за первичную обработку визуальной информации. Субъективно такие люди ничего не видят, но если попросить их высказать предположение о свойствах предмета, находящегося перед ними, то при соответствующей тренировке они дают правильные ответы. Сама возможность верных догадок связана с тем, что зрительная информация поступает в кору не только через участок V1. Но активация обходных нейронных путей не сопровождается сознанием.

На первый взгляд может показаться, что эти данные являются прекрасным подтверждением идеи эпифеноменальности сознания: сознательны мы или нет, зрение все равно работает и позволяет распознавать предметы. В действительности все как раз наоборот. Дело в том, что люди со слепым зрением — и это четко зафиксированный факт — в любом случае различают предметы гораздо хуже, чем те, у кого зрение функционирует при наличии субъективного опыта. А это означает, что субъективный опыт может влиять на распознавание.

Казалось бы это предположение можно было бы подвергнуть сомнению, если бы удалось натренировать соответствующего слепому зрению визуального «зомби-агента» (как называют подобные системы К. Кох)¹¹ до такого уровня, когда он оказался бы функционально изоморфным механизму осознанного зрения. Но на самом деле этот результат ничего не доказывал бы. Из того, что визуальная зомби-система полностью компенсировала бы утрату осознанного зрения, еще не следует, что оригинальная система функционирует без каузального участия качественных состояний. Максимум, что можно было бы сказать в этом случае, — так это то, что квалиа могут быть эпифеноменами, что их возможная каузальная роль всегда может быть продублирована чисто физическими механизмами. Забегая вперед, хочу сказать, что даже этот вывод я считаю слишком сильным. Но пока я не буду доказывать этого, а всего лишь предположу — в качестве гипотезы, — что он ложен. Одним из следствий такой гипотезы будет то, что функциональное совершенствование слепого зрения и других зомби-систем должно сопровождаться появлением или восстановлением субъективного опыта. Ряд экспериментов, похоже, подтверждает это предположение¹². Впрочем, это подтверждение не является решающим доводом в пользу вышеупомянутой гипотезы.

Суть в том, что восстановление субъективного опыта при совершенствовании зомби-систем совместимо и с эпифеноменализмом. Чтобы убедиться в этом, достаточно вспомнить основоположения натуралистического дуализма Чалмерса.

Чалмерс утверждал, что качественные состояния, лишённые каузального потенциала, скоррелированы с определёнными функциональными схемами, и из этого предположения, являющегося одной из разновидностей эпифеноменализма, точно так же вытекает, что если зомби-системы оказываются натренированными до такой степени, что они предстают в качестве функциональных изоморфов обычных систем, то их функционирование должно сопровождаться субъективным опытом.

Неустойчивое равновесие между допущением реальной действительности ментальных состояний и эпифеноменализмом, как может показаться, нарушается нашумевшими экспериментами Б. Либета. Либет показал¹³ — хотя надёжность его результатов до сих пор оспаривается¹⁴, — что мозг «принимает решение» инициировать то или иное действие ещё до того, как соответствующий волевой акт осознаётся. На первый взгляд это прямо свидетельствует в пользу эпифеноменальности наших субъективных переживаний: осознанные, качественные воления явно не успевают за реальными причинами поведения¹⁵. Однако это поспешное заключение, так как в нём содержится посылка, не являющаяся между тем очевидной. Речь идет об отождествлении качественных состояний, в том числе волений, и осознанных качественных состояний, т. е. просто качественных состояний, с одной стороны, и таких качественных состояний, о которых субъект может отчитаться как об осознаваемых им, — с другой. И хотя некоторые философы, такие, как Сёрл, полагают, что подобное отождествление правомерно, другие авторы, и прежде всего Н. Блок, отмечают, что у нас есть все основания воздерживаться от него.

Блок проводит различие между «феноменальным сознанием» и «сознанием доступа» (access consciousness)¹⁶. Наличие феноменального состояния, феноменологии, само по себе ещё не гарантирует, что оно доступно системе, называемой субъектом, т. е. не гарантирует, что оно осознаётся в собственном смысле слова. Я не буду сейчас рассматривать и оценивать аргументы Блока в пользу этого различия¹⁷. Отмечу лишь: по крайней мере мыслимо, что Блок прав, и если это так, то эксперименты Либета перестают играть на стороне эпифеноменализма. Ведь не исключено, что «решение» мозга индуцировать определенное поведение было принято при участии чисто феноменальных, т. е. качественных, но неосознанных факторов. И хотя данные факторы были осознаны уже после решения, это ещё не значит, что они не играли никакой каузальной роли¹⁸.

Таким образом, равновесие эпифеноменализма и реализма ментальной каузальности по-прежнему не нарушено. И у нас

даже может возникнуть ощущение бесперспективности подобного перетягивания каната. Но мне представляется, что это все же не так. Дело в том, что эпифеноменализм, как мне кажется, может быть опровергнут, пусть и не эмпирическими доводами, но логическими аргументами¹⁹. Надо сказать, что в последние десятилетия было предпринято немало попыток опровержения этой доктрины. Не все они были удачными, но все они заслуживают внимания. Но прежде чем перейти к разбору специальных доводов против эпифеноменализма, я постараюсь показать, что эпифеноменалистическое представление о сознании все-таки противоречит установкам здравого смысла. Совсем недавно я, правда, утверждал, что это противоречие мнимое. Но речь шла о других установках. А именно речь шла о том, что убеждение здравого смысла в каузальной действительности таких квалитативных состояний, как желания, вырастает из факта корреляции желаний и поведения. Этот факт, однако, недостаточен для вывода о причинной зависимости между ними. Таким образом, сторонник эпифеноменализма может интерпретировать установку здравого смысла именно в качестве суждения о корреляции и обвинять философских противников эпифеноменализма в неверном ее толковании.

Борьба вокруг здравого смысла весьма значима: философская система едва ли может рассчитывать на серьезное признание, если она радикально расходится с его принципами. Эпифеноменализм пока выдерживал испытание ими. Но испытание еще не закончилось. Вопрос в том, можно ли согласовать эпифеноменализм с представлением о наличии квалитативных состояний у других людей? Кажется ведь, что это представление является одной из аксиом здравого смысла. Сейчас я попробую доказать, что эти вещи несовместны.

Проблема других сознаний стала одной из самых заезженных тем современной философии, да и не только философии. Но совокупная мудрость, рожденная множеством изысканий, недалеко ушла от трюизма: я приписываю сознание другому, поскольку сам обладаю сознанием и замечаю, что другой ведет себя в чем-то похоже на меня.

Само приписывание сознания, или квалитативных состояний, — это инстинктивный процесс, реализующийся, судя по всему, через «зеркальные нейроны»²⁰. Восприятие каких-то действий других людей активирует участки мозга, ответственные за продуцирование подобных действий, — и соответствующие приватные переживания, которые, естественно, атрибутируются и тем самым людям, которые совершают эти действия. Но все это не мешает отрефлексировать логические формы этого естественного атрибутирования. Итак, в основе

лежит общий принцип: отмечая тесную связь неких состояний в одних случаях, я ожидаю ее и в других. В применении к нашей ситуации: во-первых, я отмечаю, что определенное поведение людей заставляет меня представлять свое поведение такого же рода сопровождающимся некоторыми приватными состояниями. Далее, сходство моего поведения и поведения других я считаю достаточным основанием для заключения, что где-то в приватном пространстве этих людей тоже существуют такие сознательные состояния.

Но эпифеноменалист утверждает, что поведение людей не зависит от того, наделены ли они сознанием. Сознание может исчезнуть, а поведение не изменится. Это можно понять так, что механизмы в мозге, генерирующие поведение, не зависят от механизмов, порождающих сознание. В таком случае от наличия сходного поведения нельзя заключать к наличию сознания²¹. Но именно это заключение диктует нам здравый смысл. Поэтому эпифеноменализм противоречит установкам здравого смысла, что и требовалось доказать.

Впрочем, эпифеноменалист может заявить, что его позиция совместима с тезисом, согласно которому механизмы, генерирующие поведение, являются именно теми механизмами, которые порождают сознание. Ведь даже если эти механизмы совпадают, сознание все равно может быть эпифеноменальным, т. е. не оказывать влияния на работу мозга. Такой взгляд, напомню, отстаивает Чалмерс. Он нигде не утверждает, что мозг может реально функционировать — т. е. порождать соответствующее поведение — при отсутствии сознания. Конечно, подобная ситуация мыслима, но не более того. В реальности же везде, где есть нормально работающий мозг, есть и порождаемое им сознание. Поэтому от сходного с моим поведения я могу заключать к существованию сознания, даже и будучи эпифеноменалистом. Подобные возражения кажутся резонными, но вопрос в том, откуда эпифеноменалист знает, что механизмы генерации поведения совпадают с механизмами порождения сознания.

В самом деле, даже если согласиться, что механизмы генерации поведения участвуют в порождении сознания, почему бы не допустить, что в нем участвует и некий дополнительный фактор? Если бы он не был задействован, все процессы в мозге протекали бы «в темноте», генерируя при этом одинаковое с моим поведение. Хочу напомнить, что Чалмерс рассматривает такую возможность и в предварительном плане отвергает ее по следующей причине: такое допущение нарушает эпистемологический принцип наиболее простого объяснения. Он не вдается в подробный анализ своего утверждения, полагая, что связь паттернов в мозге и сознания можно дока-

зять с использованием более сильных доводов, а именно с помощью мысленных экспериментов «блекнущие квалиа» и «скачущие квалиа». Мы уже знаем, однако, что эти эксперименты не достигают цели, и поэтому вся тяжесть ложится на рассуждение о том, что в противном случае мы без необходимости усложняем теорию.

Впрочем, это рассуждение выглядит вполне правдоподобным. И впрямь на первый взгляд гораздо экономнее допустить, что сознание связано исключительно с функциональными схемами в мозге (т. е. с механизмами, генерирующими поведение организма), чем с независимым фактором или с функциональными схемами и плюс к ним с каким-то неизвестным фактором. Но это лишь на первый взгляд. С этим положением можно было бы согласиться, если бы не одно «но». И «но» это состоит в том, что признание связи сознания с одними лишь функциональными схемами приводит к еще большему усложнению мира, чем допущение какого-то дополнительного X-фактора. Ведь в таком случае получалось бы, что мы действительно должны допускать существование эпифеноменальных сознательных состояний у людей и других существ, демонстрирующих сходные со мной паттерны поведения. Но так как подобные эпифеноменальные состояния, по определению, не играют роли в детерминации поведения этих существ, то мы буквально наводнили бы мир состояниями, совершенно ненужными для предсказания его изменений во времени. Это было бы классическим умножением сущностей безо всякой на то необходимости. Соответственно, экономнее допустить, что сознание порождается не функциональными схемами в мозге или не только ими, а возникает при помощи X-фактора. Однако эта более экономная гипотеза блокирует вывод о наличии у других людей сознательных состояний — до тех пор, пока мы не сможем достоверно сказать, что это за фактор и присутствует ли он у них. Таким образом, если здравый смысл заставляет нас верить в существование сознательных состояний у других людей, этот же здравый смысл принуждает нас отказаться от эпифеноменализма.

Антиэпифеномалистическая позиция подразумевает, что сознание влияет на поведение. И хотя само по себе это еще не является основанием для утверждения, что сходное с моим поведение сопровождается сознанием, поскольку одинаковые действия могут порождаться разными причинами, и, абстрактно говоря, можно представить, что чисто механическая система будет продуцировать поведение, неотличимое от моего, допущение подобного разнообразия причин на этот раз действительно будет вступать в противоречие с тем самым принципом простоты, к которому безуспешно апеллировал эпифе-

номеналист²². Впрочем, чуть позже в этой самой главе я попытаюсь показать, что механические системы в принципе не могут дублировать человеческое поведение, — и если это удастся сделать, то наличие сходного с моим поведения можно будет рассматривать в качестве верного признака присутствия другого сознания.

Однако не будем забегать вперед и попробуем все же защитить эпифеноменализм — насколько это вообще возможно. Л оборонительные ресурсы этой доктрины еще не исчерпаны. Ведь эпифеноменалист может сказать, что все это заигрывание со здравым смыслом — очень рискованное предприятие. Может быть, здравый смысл и впрямь идет против эпифеноменализма — но насколько он надежный провожатый? Какие только нелепости не порождает этот здравый смысл! Он очень грубый инструмент. Да и вообще, что такое здравый смысл? Можно ли всерьез работать с его понятием?

Это важные вопросы, и вскоре мы займемся ими. Мы увидим, что попытка прояснения понятия и аксиом здравого смысла приведет нас к выводу, что он может играть и на стороне эпифеноменализма, правда лишь в том случае, если это прояснение неполно. Пока же пойдем на поводу у эпифеноменалиста, рассуждающего о расплывчатости понятия здравого смысла, и посмотрим, нельзя ли доказать ошибочность эпифеноменалистской установки без обращения к этому понятию. Я уже говорил, что в последнее время был предпринят ряд интересных попыток подобного доказательства. Начнем, пожалуй, с попытки, которая в действительности является лишь новым изданием довольно старого аргумента, разработанного еще У. Джемсом в ответ на эпифеноменалистские построения Гексли.

Речь идет о так называемом эволюционном аргументе. Он очень прост. Если мы признаем, что человек, как и другие животные, возник под давлением естественного отбора, то логично предположить, что базовые характеристики человека как биологического вида имеют определенную приспособительную ценность. Но эпифеноменалисты отрицают всякую подобную ценность за сознанием, так как оно не оказывает никакого влияния на адаптивное поведение. Соответственно, эпифеноменалист не может объяснить, почему сознание вообще эволюционировало. И наоборот, если мы признаем, что сознание — это продукт эволюции, мы должны соглашаться с тем, что оно оказывает полезное влияние на человеческое поведение.

Несмотря на наглядность этого рассуждения, надо сказать, что оно не пользуется большой популярностью у современных авторов. И связано это с тем, что оно играет на стороне интеракционистского дуализма, который не очень привле-

кателен для натуралистически настроенных — в основной своей массе — аналитических философов. Неудивительно поэтому, что среди защитников этого аргумента в наши дни встречаются мыслители, тяготеющие к теологическим темам, — такие, как У. Хескер. В книге «Эмерджентная самость» (1999) он уверяет, что размышления над проблемой эволюции сознания (которые, по его мнению, могут быть направлены не только против эпифеноменализма, но и против физикализма) приводят нас к следующему выводу: «надо отказаться от каузальной замкнутости физического, а значит, и от приверженности универсальным механическим объяснениям. Сознательные ментальные состояния должны быть признаны каузально эффективными именно благодаря их ментальному содержанию. Если признать это, то данные состояния могут влиять на поведение и быть подверженными давлению отбора»²³.

Но так ли уж силен эволюционный аргумент? Ведь эволюционная теория не утверждает, что всякая характеристика организма имеет приспособительное значение. Т. е. она, как уже отмечалось, не накладывает запрета на существование нейтральных в приспособительном плане или даже негативных характеристик, если только они являются побочными

продуктами позитивных свойств²⁴, появление которых с избытком компенсирует минусы этих побочных продуктов.

Иными словами, существование эпифеноменального сознания по крайней мере не противоречит основам эволюционной теории²⁵. Нельзя, правда, говорить, что теория эволюции объясняет появление эпифеноменального сознания, но отсутствие объяснения и противоречие — разные вещи. Хескер, впрочем, уверен, что отсутствие объяснения — тоже очень плохо. Соответствие сознательных состояний и процессов в мозге становится чем-то вроде чуда. Но и на это можно возразить, сказав, как это делает Чалмерс в «Сознающем уме», что речь здесь идет о фундаментальном законе, а фундаментальные законы не допускают дальнейшего объяснения²⁶. Что можно на это ответить? Разве что спросить, в свою очередь, Чалмерса, откуда ему известно об этом фундаментальном законе и его предпосылках — принципе когерентности и принципе организационной инвариантности? На этот вопрос Чалмерс, как я уже не раз отмечал, не в силах дать удовлетворительного ответа, но важно, что сам этот вопрос, раскрывающий слабое место позиции Чалмерса, не имеет прямого отношения к эволюционному аргументу. Наоборот, эволюционный аргумент мог бы обрести хоть какую-то видимость убедительности лишь после независимого обнаружения дефектов в позиции Чалмерса.

Я не случайно все время возвращаюсь к Чалмерсу. Ведь именно он взял на себя труд систематизации возражений против эпифеноменализма и демонстрации их очевидной недостаточности. Помимо эволюционного аргумента и одного из возражений с позиции здравого смысла (а именно первого из рассмотренных в этой главе) в «Сознающем уме» он обсуждает и еще один довод, который кажется ему наиболее опасным и который он анализировал еще в рукописи 1990 г. Я имею в виду аргумент, права на авторство которого отстаивал израильский физик и философ Авшалом Элитцур. Мы уже встречались с этим аргументом. Если эпифеноменализм верен, то наше поведение не зависит от сознания. Не зависит от сознания, стало быть, и вербальное поведение, в том числе высказывания о самом сознании, его отличии от мозга, о загадке сознания и т. п. Это кажется парадоксальным. Получается, что зомби — существо, лишенное сознания, — с не меньшим основанием могло бы говорить о «трудной проблеме сознания» и пытаться решить ее.

Чалмерс подробно рассматривает это возражение, именуемое им «парадоксом феноменального суждения». Его вывод заключается в том, что надо смириться с данным парадоксом и научиться жить с ним. Т. е. он не считает, что он приводит к разрушению эпифеноменализма. Но все это, конечно, заслуживает более подробного анализа.

Важно прежде всего понять, в каком случае упомянутый парадокс может привести к разрушению эпифеноменализма. Чалмерс видит опасность в следующем вопросе: убеждены ли мы в существовании сознания? Суть в том, что убеждения в аналитической философии традиционно понимаются как такие состояния, которые можно истолковать в терминах поведенческих диспозиций. К примеру, если я убежден, что в комнате компьютерный монитор, то это значит, что я predisposed смотреть на его экран, включать его в начале работы и т. п. Но поведение и соответствующие диспозиции, согласно эпифеноменализму, не зависят от сознания. Получается, что само наше убеждение в существовании сознания никак не связано с его существованием. Это довольно тревожная ситуация, так как обычно если убеждение в чем-то не связано (как правило, каузально) с существованием того, в чем мы убеждены, то такое убеждение является ложным или по крайней мере лишенным реального основания (в случае с упоминавшимся выше монитором, если мое убеждение, что он находится в комнате, не связано с тем, что он действительно находится в комнате и вызывает у меня соответствующие ощущения, то тогда это галлюцинация). Одним словом, если наше убеждение в существовании сознания каузально не свя-

зано с сознанием, то у нас нет оснований допускать существование сознания как чего-то отдельного от поведенческих диспозиций. На это Чалмерс возражает, что не всякое убеждение можно понять через предрасположенность к какому-то поведению. И некоторые истинные убеждения могут быть наделены такими содержаниями, которые даются не в результате каузального влияния, а, так сказать, непосредственно. И это возражение²⁷ сохраняет status quo. Хотя все, что я говорю о сознании, зависит от физических процессов и, возможно, допускает толкование в физических или функциональных терминах, я непосредственно знаю, что мои слова имеют и другой, качественный смысл. И такое предположение трудно опровергнуть.

Впрочем, Роберт Кирк - этот Франкенштейн аналитической философии — в книге «Зомби и сознание» (2005) попробовал сделать это²⁸. Эпифеноменалисты чалмерсовского толка, замечает он, признают, что квалиа не имеют каузального влияния на физические процессы, но в то же время не отрицают, что мы можем думать о квалиа, обращать на них внимание и т. п. Правда, утверждают они, это происходит не потому, что квалиа реально действенны, а в силу того, что соответствующие когнитивные процессы находятся в отношении «эпистемической близости» с квалиа — она и есть то, что Чалмерс называет непосредственной осведомленностью о качественных состояниях. Так вот, Кирк пытается показать, что подобная близость немыслима и невозможна. Для этого он придумывает на редкость странный эксперимент. Представим зомби, моего принципиального двойника, но такого, что процессы обработки визуальной информации, протекающие в его мозге, порождают цветные картинки на его ступнях, причем эти картинки (о которых он ничего не знает, иначе они повлияли бы на его поведение, что специально исключается в данном случае) соответствуют тем визуальным квалиа, которые были бы у меня, его двойника, в аналогичной перцептивной ситуации. И Кирк задается вопросом: есть ли у нас хоть какие-то основания считать, что это кино на ступнях зомби находится в отношении эпистемической близости к процессам в его мозге? Очевидно, что нет. Но чем, в сущности, отличаются мои эпифеноменальные квалиа от такого фильма? Кирк считает, что ничем, и делает отсюда вывод, что и квалиа эпифеноменалиста не могут состоять в отношении эпистемической близости к мозговой деятельности.

Предполагаемая непредставимость сценария эпистемической близости, а значит, и эпифеноменализма подталкивает Кирка к выводу о непредставимости зомби, так как представимость зомби, утверждает он, имплицитно предстает-

мости рассматриваемого и опровергаемого им варианта эпифеноменализма. При этом он странным образом забывает, что сама его критика эпифеноменализма опирается на мысленный эксперимент, в котором фигурирует зомби²⁹. Но даже если отвлечься от зомби (в скобках замечу, что никак не могу понять причины увлеченности многих проблемой представимости зомби³⁰, здесь нет никакой проблемы: другие люди непосредственно воспринимаются нами именно в качестве зомби и проблема состоит в другом — в отыскании весомых оснований для того, чтобы не считать других людей зомби³¹) и сконцентрироваться на критике Кирком эпифеноменализма, то окажется, что предложенный им мысленный эксперимент, по-видимому, несложно нейтрализовать.

Задумаемся еще раз: почему Кирк считает, что между картинками на ступнях и процессами в мозге зомби нет эпи-стемической близости? Потому что зомби не замечает их, хотя они порождаются им. А так как между сменяющимися картинками на ступнях зомби и эпифеноменальными квалиа нет принципиальной разницы, то эпифеноменальные квалиа тоже не могут быть замечены. Поэтому между когнитивными процессами в мозге, связанными с отчетами о квалиа, размышлениями о них и т. п., и самими квалиа не может быть эпистемической близости, которая требуется эпифеноменалисту для разрешения чалмерсовского «парадокса феноменального суждения».

Чем больше прокручиваешь этот сценарий, тем менее убедительным он кажется. Кирк считает, что доказал, что если квалиа эпифеноменальны, то они не могут быть замечены. Но кем замечены? Если телом, то эпифеноменалист с этим не спорит: тело реагирует только на физические процессы. Если Я, то что такое Я помимо квалиа? Суть в том, что поскольку квалиа составляют сознание, то качественные состояния можно мыслить (хотя это не обязательно) как синоним осознанных, а значит, кем-то замеченных состояний. И раз тело не мол-сет заметить квалиа, они замечаются непосредственно.

Это именно то, о чем говорит Чалмерс и с чем несогласен Кирк. Но теперь мы видим, в чем слабость его аргумента: он неоправданно преуменьшает различие между картинками на ногах зомби и квалиа. Различие состоит именно в том, что в понятии квалиа уже присутствует или по крайней мере может присутствовать замеченность. И это обстоятельство, похоже, разрушает всю логику его рассуждений. Кирк пытается сказать, что при эпифеноменалистских предпосылках квалиа не могут быть замечены, что, помимо прочего, как утверждает он, противоречит фактам. В действительности это может противоречить самому определению квалиа, но такое проти-

воречие уже не играет на его стороне, так как получается, что, даже если квалиа эпифеноменальны, они все равно замечаются. А это, в свою очередь, означает представимость того самого сценария, невозможность которого он хотел показать, — и все остается на своих местах³².

Подведем промежуточные итоги. Недавние аргументы против эпифеноменализма, не опирающиеся на интуиции здравого смысла в качестве основы для критических выпадов против этой доктрины, оказались не столь перспективными, как можно было бы подумать. С другой стороны, предложенный выше аргумент, исходящий из аксиом здравого смысла, наталкивается на возражение о непроясненности самого этого понятия. И очевидно, что нашим следующим шагом должно стать уточнение природы интуиции здравого смысла.

К вопросу о здоровом смысле можно подойти по-разному. Можно раскрывать это понятие, так сказать, из него самого, как это, к примеру, делал Дж. Э. Мур³³. Но тут, конечно же, возникает опасность круга. Можно стартовать с истории этого понятия. А история эта началась еще в Античности, но кульминация пришлась на XVIII век. С подачи Э. Шефтсбери это понятие проникло сначала в британскую философию, а затем распространилось по всей Европе. Но главными экспертами в области здравого смысла оставались шотландцы, представители так называемой школы «здравого смысла». Один из них, Дж. Битти, проанализировав ошибочные толкования «здравого смысла» (*common sense*), отождествляющие его с благоразумием, вкусом, т. е. с «общим чувством» (*common sense*), а также с пронизательностью и общим (*common*) мнением, провозгласил, что здравый смысл — это «способность ума, постигающая истину или направляющая веру — не путем последовательной аргументации, но мгновенным, инстинктивным, непреодолимым импульсом, полученная не в результате образования, не из привычки, но от природы, действующая независимо от нашей воли при наличии объекта по установленным законам и поэтому справедливо называемая интуицией (*sense*) и воздействующая сходным образом на всех или, по крайней мере, на подавляющее большинство людей»³⁴. Битти, правда, противопоставлял здравый смысл разуму как способности умозаключать, исходя из принципов здравого смысла, но другие философы этой школы, такие, как Т. Рид и его ученик Д. Стюарт, настаивали, что здравый смысл — это интуитивный аспект самого разума³⁵.

Так или иначе, в трактовке здравого смысла британскими авторами эпохи Просвещения принципиально важными являются две вещи. Во-первых, здравый смысл есть резервуар истин, с которыми не может не согласиться ни один человек,

т. е. универсальных истин. Во-вторых, эта универсальность несинонимична строгой необходимости: к числу истин здравого смысла относятся и такие, которые в известном смысле случайны, т. е. отражают такое положение вещей, которое могло бы и не существовать — и противоположное им мыслимо. Поскольку, однако, истины такого рода тем не менее универсальны, их универсальность можно объяснить тем, что они возникают, как говорили, «из человеческой природы»³⁶, т. е. необходимо связаны с устройством познавательных и иных способностей человека.

Сразу хочу сказать, что такие объяснения истоков здравого смысла кажутся мне вполне убедительными, и в этой книге я присоединяюсь к ним. При этом под принципами здравого смысла я в первую очередь буду понимать именно положения, необходимо связанные с устройством наших познавательных способностей. Таких положений, разумеется, может быть много, но не исключено, что какие-то из них можно будет редуцировать к другим. Поиском основоположений здравого смысла нам сейчас и предстоит заняться. И прежде всего надо конкретизировать понятие познавательных способностей, из устройства которых, согласно предположению, должны вытекать принципы здравого смысла. Если, к примеру, согласиться со сторонниками модульной теории сознания и эволюционной психологии, то окажется, что познавательные способности есть не что иное, как когнитивные модули в мозге, играющие определенную адаптивную роль. И если это так, то изучать познавательные способности можно исключительно эмпирически — философия с ее концептуальным анализом попросту лишается работы. Но, на мой взгляд, какая-то часть работы все же может быть проведена и априори.

Все дело в том, что в человеческом познании, поскольку оно нацелено на факты, а не на логические соотношения, есть одна подлинно универсальная черта, а именно: все это познание построено на переносе прошлого опыта на будущее, как прекрасно показал еще Юм.

Действительно, любое суждение, которое мы выносим о вещах — постольку, поскольку высказанное в них не отражает то, что непосредственно дано в чувствах, а выходит за пределы наличного, — подразумевает уверенность в правомерности определенных ожиданий. Скажем, если я говорю: «Это хрупкая вещь!», это значит, что я ожидаю, что в определенной ситуации с ней произойдет то-то и то-то, а именно что она разобьется. Можно взять наугад любой пример суждения о фактах (но, повторю, только такого, где дело не ограничивается описанием непосредственно данного), и в нем обнаружится то же самое. Итак, очевидно, что в подобных (синтетических) сужде-

ниях о фактах неявно утверждается нечто о будущем. Но априори, как опять же давно известно, нельзя достоверно, сконструировать будущее, так как подобное конструирование возможно лишь в ситуации, где нет альтернатив, а будущее всегда можно мыслить альтернативно. Значит, наши суждения о будущем базируются не на априорных размышлениях, а на переносе на будущее уже имеющейся у нас — т.е. почерпнутой в прошлом опыте — информации.

Таким образом, если мы убеждены в чем-то касающемся фактов или попросту вещей и их свойств, то мы должны быть убеждены и в правомерности переноса прошлого опыта на будущее. Т. е. мы должны быть убеждены в соответствии прошлого и будущего.

Откуда берется это убеждение? Речь в нем не идет о чем-то самоочевидном, да и из опыта нельзя генерализировать тезис «будущее соответствует прошлому», так как подобное обобщение возможно лишь при предположении соответствия прошлого и будущего.

Юм, который впервые произвел анализ веры в соответствие прошлого и будущего, полагал, что она проистекает из привычки как одного из врожденных принципов человеческой природы. Иногда он высказывался в таком ключе, будто этот принцип — состоящий именно в склонности переносить прошлый опыт на будущее — имеет характер инстинкта. Подобное понимание выглядит вполне современно и хорошо вписывается в контексты эволюционной психологии.

Не исключено, однако, что, рассуждая о склонности переносить прошлое на будущее как особом когнитивном инстинкте, Юм все же несколько переусложнил картину. Можно попробовать показать, что убеждение в соответствии прошлого и будущего с неизбежностью проистекает из самого устройства человеческого сознания. Но прежде чем мы сделаем это, хочу предупредить читателей, что следующие страницы могут вызвать некоторые затруднения, так как все доводы будут разворачиваться в феноменологической плоскости, и, более того, сама эта феноменологическая плоскость будет не совсем обычной для феноменологических работ, так как это будет не дескриптивная феноменология, характерная, к примеру, для Гуссерля, а аргументативная, или аналитическая, феноменология, с давних пор привлекавшая мое внимание³⁷.

Итак, для прояснения механизмов формирования убеждения о соответствии прошлого и будущего примем вначале, что все представления, которые мы можем составить о будущем, состоят из компонентов прежних ощущений. Мы не в состоянии отчетливо представить принципиально новый вид будущих ощущений. Для иллюстрации этой невозможности

допустим, что нам станет достоверно известно, что через какое-то время у нас откроется новое чувство, к примеру позволяющее нам воспринимать ультрафиолетовое излучение. Как будут выглядеть наши новые ощущения, мы не в силах вообразить. Если согласиться с этим общепринятым с времен Локка тезисом, то это и будет означать, что все наши проекты будущего составлены из комбинаций прежних ощущений. Любой компонент этих гипотетических сочетаний можно, стало быть, трактовать как воспоминание или по крайней мере как воспроизведение. Но всякое воспроизведение имеет темпоральный характер. Иными словами, воспроизвести что-то — значит репродуцировать не отдельное атомарное представление (это вообще, видимо, некорректное понятие), а ряд представлений, в «середине» которого оно оказывается³⁸.

Далее, очевидно, что онтологический статус воспроизведенных компонентов ряда не может отличаться от онтологического статуса «центрального» представления — ведь в прошлом, из которого они извлекаются, все они были однородны. Т. е. если, к примеру, мы наделяем воспроизведенное представление, используемое нами для конструирования картины будущего, статусом «квазиощущения» (это происходит, когда мы верим, что произойдет событие, обозначаемое данным представлением), то квазиощущениями должны оказываться и компоненты темпорального ряда, составлявшие в прошлом опыте — или, точнее, в том, что память преподносит в качестве такового³⁹ — «окрестность» данного представления, распространяющуюся как «вперед», так и «назад», т. е. как в прошлое, так и в будущее относительно ожидаемого события. Соответственно, какой-то из компонентов этого воспроизведенного ряда должен позиционироваться в момент «теперь», относительно которого у нас имеется то или иное наличное ощущение. Если это воспроизведенное квазиощущение не тождественно наличному ощущению, то возникает противоречие, так как получается, что мы верим в наличие того, в наличие чего мы как раз не верим. Весь ряд в таком случае лишается статуса квазиощущений. Мы не можем верить в то, что произойдет событие, входящее в ряд воспоминаний, не отсылающий к тому, что ощущается в данный момент. Т. е., мы можем верить только в ряд событий, идентичный тому, в который было включено наличное ощущение (всегда служащее для нас отправной точкой отсчета при размышлении о будущем) в прошлом. Это и значит, что мы не можем не верить в соответствие прошлого и будущего опыта, причем в их строгое соответствие, тождество⁴⁰. Можно сразу возразить, что в действительности мы очень часто верим в то, что произой-

дут события, не входящие в точности в ряды, имевшие место в прошлом. Но это возражение ничего не отрицает. Мы/ведь рассматривали идеальный случай. Если ряды не в точности совпадают, то вера просто ослабляется. Градаций этой веры может быть очень много. Но само ослабление веры при неточном повторении ее прежних условий подтверждает, что при их точном повторении она была бы максимальной.

Я хочу подчеркнуть и развить этот момент. Признание фундаментальности тезиса «я верю в тождество прошлого и будущего» не означает, что его можно истолковать так, будто мы должны верить, что будущее повторяет прошлое. Мы вовсе не обязаны считать так, скорее, мы думаем наоборот. Но это не противоречит сказанному выше. Смысл веры в тождество прошлого и будущего может быть выражен следующим образом: мы верим, что при точном повторении в будущем тех событий, которые имели место в прошлом, за ними последует то, что следовало за ними тогда, когда мы раньше ощущали их. Из этой формулы вовсе не вытекает, что такие точно повторяющиеся события действительно будут иметь место. Необходимости здесь нет: одно состояние может влечь за собой другое, отличное от него, оно — третье, опять-таки отличное, и так до бесконечности.

Несколько формализуя сказанное об убеждении в тождестве прошлого и будущего, получаем, что мы верим, что при повторении всех компонентов прошлого события (включая его композицию, или структуру) за ним последует то же совокупное событие⁴¹, что и раньше. Возьмем теперь какой-нибудь из компонентов этого будущего совокупного события, т. е. любое «событие» в «Событии»⁴². Поразмыслив, мы заметим, что наша аксиома тождества приводит к любопытным выводам по поводу того, что мы должны думать (точнее, во что мы должны верить) относительно данного сингулярного события. Указанное событие должно как-то соотноситься с прошлыми событиями, коль скоро совокупное событие, в которое оно входит, определенным образом соотносится нами с тем, что предшествует ему. С другой стороны, очевидно, что многие сингулярные события повторяются в совершенно другом событийном окружении. Так что приходится допустить — я назову это неизбежное для нас, но эмпирическое по сути допущение тезисом об автономности и локальности событийных рядов, — что любой произвольно выбранный элемент будущего события может связываться нами лишь с некоторыми из компонентов нынешнего совокупного события. Причем эта связь должна мыслиться по модели жесткого следования.

Последнее предположение может быть подтверждено строгим доказательством. Допустим, что произвольно выбранный

нами компонент ожидаемого будущего События не имеет необходимого коррелята в Событии, предшествующем ему. Необходимая корреляция в данном случае означает, что одно событие постоянно сопровождается другим. Если у выбранного сингулярного события нет подобного коррелята в предшествующем Событии, т. е. если в составе предшествующего События нет события, наличие которого всегда сопровождается нашим событием, то это равносильно тому, что предшествующее совокупное событие может повторяться без того, чтобы в последующем Событии встречалось то сингулярное событие, о котором идет речь. Но тогда это последующее Событие реально может отличаться от того, как оно выглядело в прошлом, когда ему предшествовало другое Событие, ныне в точности воспроизведенное, и, соответственно, в эту возможность вполне можно поверить. Иными словами, допущение того, что некий компонент какого-либо совокупного события не имеет необходимого коррелята в предшествующем Событии, равносильно отрицанию необходимости веры в тождество прошлого и будущего в указанном смысле. Соответственно, признание тождества прошлого и будущего влечет за собой вывод о наличии необходимых коррелятов в предшествующем Событии у любого из компонентов последующего совокупного события.

Поскольку вера в тождество прошлого и будущего необходимо интегрирована в саму структуру сознания, она неизбежна для нас. Столь же неизбежной, стало быть, оказывается и вера в необходимую корреляцию с прошлым отдельных компонентов совокупных событий. Но подобная корреляция составляет самую суть каузальных, причинностных отношений. Согласно одной из эталонных дефиниций, причина есть событие А, так связанное с неким событием В, что наличие А всегда приводит к появлению В. Нетрудно заметить, что это определение полностью соответствует разобранному случаю. Вера в тождество прошлого и будущего, мы одновременно верим, что любому сингулярному событию соответствует что-то в предыдущем совокупном событии, что соотносено с ним таким образом, что его появление всегда приводит к тому, что это сингулярное событие происходит. В общем, из веры в тождество прошлого и будущего вытекает вера в то, что каждое событие имеет причину.

Следует отметить, что аксиома тождества не порождает веру в то, что каждое сингулярное событие должно иметь только одну причину. Такая вера действительно возникает, если речь идет о совокупном событии. Некое Событие А, как мы верим, может сопровождаться только вполне определенным Событием В. Иначе происходило бы нарушение веры

в тождество прошлого и будущего. Обозначенный выше механизм порождения этой веры работает таким образом, что альтернатив здесь просто нет. Но выводы, сделанные относительно совокупных событий, не распространяются на их сингулярные компоненты. В самом деле, повторение какого-либо компонента с другим коррелятом в прошлом вполне допустимо в контексте других Событий. Но какие-то необходимые корреляты быть должны, и именно это самое важное в структуре каузальной веры.

Теперь можно на мгновение остановиться и посмотреть, что же мы, собственно, достигли в результате этих умозаключений. Мы занимались понятием здравого смысла в его когнитивном аспекте и выяснили, что к истинам здравого смысла можно причислить универсальные положения, необходимо связанные с механизмом человеческого познания. И мы нашли такое, положение - принцип тождества прошлого и будущего. Заметим, правда, что, как таковое, это положение не является тезисом здравого смысла. Скорее, его надо выявлять с помощью рефлексии. Но это не значит, что оно не имеет отношения к когнитивному здравому смыслу. Ведь оно может быть неким основоположением, неявным истоком других положений, принадлежность которых к здравому смыслу не вызывает сомнений.

Так оно и оказалось. Как мы только что видели, аксиома тождества, или соответствия, лежит в основе убеждения в том, что каждое событие имеет причину. А это убеждение, конечно же, является одной из важнейших установок здравого смысла. Существенно при этом, что демонстрация связи каузальной веры и аксиомы тождества позволяет уточнить природу первой. Дело в том, что сама механика переноса прошлого опыта на будущее такова, что она заставляет не просто верить, что каждое событие имеет причину, но и размещать эти причины среди данностей наших чувств, среди предметов опыта. И это обстоятельство возвращает нас к проблеме ментальной каузальности и эпифеноменализма. Посмотрим, как это происходит.

Итак, я с необходимостью верю, что каждое событие имеет причину в опыте. Мой опыт многообразен и включает в себя самые разные данности, в том числе приватные аспекты таких ментальных состояний, как желания, эмоции и т. п. Поскольку опыт также показывает, что эти качественные аспекты скоррелированы с другими данностями, в частности с теми, которые могут быть охарактеризованы как поведенческие, то у меня есть все основания считать, что квалиа могут быть причинами моего поведения. Но тут возникает трудность. Ведь в опыте мне даны и другие люди, причем они даны

мне так, что я никогда непосредственно не воспринимаю их качественные состояния.

Иными словами, качественные состояния других людей не являются опытными данностями (даже если допустить, что мы смогли бы соединять мозги в единую нейронную сеть⁴³, это еще не означало бы, что я испытывал бы то, что чувствуют другие: данностями в лучшем случае были бы либо бесплодные дубликаты их частных состояний⁴⁴, либо — при более тесной интеграции — мне могли бы открываться сами состояния такого рода, но уже не других, а мои собственные, что противоречит условиям). Вместе с тем я убежден, что поведение других людей имеет определенные причины, и, более того, как было только что отмечено, я убежден, что они должны быть даны в моем опыте. Поскольку качественные состояния других людей (если они вообще существуют) не даны мне в опыте, я должен придерживаться убеждения, что их поведение детерминируется иными причинами, а именно какими-то физическими факторами.

Это же убеждение распространяется не только на поведение людей, но и на другие физические процессы. Иными словами, само устройство моей когнитивной способности навязывает мне тезис о каузальной замкнутости физического. Может, правда, показаться, что я могу оставить лазейку для самого себя, но многие сочли бы, что признание собственной уникальности оказывается в такой ситуации тупиковым решением. Экономнее распространить принцип каузальной замкнутости физического и на самого себя, объявив мнимые частные причины моего поведения эпифеноменами.

Задержимся и на этом моменте. В аналитической философии бытовало мнение, что принцип каузальной замкнутости физического, т. е. положение, что каждое физическое событие имеет физическую причину⁴⁵, связано с физической наукой Нового времени, будучи предпосылкой последней⁴⁶. Однако при ближайшем рассмотрении оказывалось, что подобное мнение лишено серьезной фактологической базы, и Д. Папино в книге «Размышляя о сознании» (2002) развенчал этот миф, невольно приоткрывая дверь новому интеракционизму⁴⁷. Прделанный выше анализ, однако, грозит захлопнуть ее. Ведь если приведенные доводы верны, то тезис о каузальной замкнутости физического имеет отношение не к науке Нового времени — равно как и не к «постулатам материалистической метафизики»⁴⁸, - а к универсальным когнитивным механизмам сознания. У нас просто нет альтернативы вере в каузальную замкнутость. Но кажется, что в таком случае мы обречены быть эпифеноменалистами.

Теперь самое время вспомнить, что ранее мы пришли к совершенно другим выводам. Разбирая соотношение эпифеноменалистских установок и положений здравого смысла, мы выяснили, что они несовместимы. Если мы стоим на позиции эпифеноменализма, то не можем придерживаться убеждения (диктуемого здравым смыслом), будто другие люди обладают сознанием. Теперь же у нас получилось, что, если мы хотим придерживаться установок здравого смысла, мы должны допускать эпифеноменальность сознания.

Здесь какое-то удивительное противоречие. Рассмотрим его получше. Еще раз вспомним, как формируется убеждение в существовании сознательных состояний у других людей. Я знаю, что обладаю подобными состояниями, вижу, что они скоррелированы с моим поведением, замечаю сходное поведение у других людей и заключаю, что у них тоже есть аналогичные качественные корреляты. Ранее я пытался показать, что если мы придерживаемся эпифеноменалистских взглядов, то лишаемся права делать подобное заключение, так как эпифеноменализм разрывает связь между поведением и сознанием. А теперь мы увидели, что эпифеноменалистский взгляд на сознание вроде бы диктуется когнитивным принципом тождества прошлого и будущего. Если так, то, дабы избежать нестыковок (тем более опасных, что принцип аналогии, использующийся при заключении о существовании других сознаний и подталкивающий к отрицанию эпифеноменализма, произведен от принципа тождества прошлого и будущего, который, как кажется, ведет нас к эпифеноменализму), приходится — раз не остается ничего другого — дисквалифицировать убеждение в существовании сознания у других людей и перестать рассматривать его в качестве надежного принципа здравого смысла, оставив ему малопривлекательную роль когнитивной иллюзии, возникающей прежде всего вследствие грубости мышления — ведь в таком случае она основана исключительно на грубой аналогии.

Но хотя такой вариант можно рассматривать, он едва ли вызовет восторги. В самом деле, еще раз оценим, какая же удивительная теория сознания у нас получается: (1) поскольку я верю в тождество прошлого и будущего, я верю, что каждое событие имеет причину, которая может быть дана мне в опыте; (2) события, происходящие с другими людьми, тоже должны иметь опытные причины; (3) мне не могут быть даны в опыте сознательные состояния других людей; (4) значит, их поведение детерминировано теми аспектами их личности, которые даны мне в опыте, а именно физическими факторами; (5) будучи человеком, я тоже должен быть детерминирован в своем поведении физическими причинами; (6) значит, мои

сознательные состояния не играют никакой роли в определении моего поведения, являясь чистейшими эпифеноменами; (7) я считаю, что другие люди обладают эпифеноменальными сознательными состояниями, хотя у меня нет весомых оснований приписывать их им.

Обратим внимание на некоторые дефекты этой теории, которая, кстати говоря, никак не решает нашу «трудную проблему». Одним из ее слабых звеньев являются пункты 5 и 6. Здесь я заключаю об эпифеноменальности собственного сознания, исходя из того, что тоже являюсь человеком, а у других людей поведение детерминировано физическими причинами. Это заключение было бы более или менее убедительным, если бы я мог сделать вывод, что другие люди, как и я, обладают сознательными состояниями. Но эпифеномалист-ская позиция дисквалифицирует такой вывод. Если же мне неизвестно, обладают ли другие люди сознанием, то у меня нет оснований приравнять себя к ним. В конце концов в собственном опыте я вижу корреляцию сознательных состояний и поведения, и эти сознательные состояния могут быть его причинами. В таком случае не менее логично считать, что другие люди — это просто зомби, а вот мое поведение детерминировано в том числе и моим сознанием. Напомню теперь, что эта «улучшенная» теория, как и ее исходный вариант, являются не чем иным, как выводами из основоположения о тождестве прошлого и будущего, лежащего в основе установок здравого смысла. Но, думаю, никто не будет спорить, что нет ничего более далекого от здравого смысла, чем эта теория. И столь громадное расхождение свидетельствует о том, что где-то в процессе рассуждений мы сбились с пути. И весь вопрос в том, где это произошло?

Прежде чем ответить на этот вопрос, вернемся к теме опровержения эпифеномализма, которая ведет эту главу, и отметим, что на той стадии, на которой мы сейчас находимся, его судьба по-прежнему остается туманной. С одной стороны, аргумент против эпифеномализма с позиций здравого смысла, который я выдвинул в начале главы, значительно ослаблен в результате осознания нами возможных эпифеномена-листских следствий фундаментального когнитивного убеждения в тождестве прошлого и будущего. С другой стороны, эти следствия не являются безальтернативными, так что эпифе-номализм тоже не получил решающего перевеса.

Остается лишь надеяться, что после того, как мы поймем, где произошел сбой в предшествующем анализе, мы сможем сказать о судьбе эпифеномализма что-то более определенное. Сбой же, как мне кажется, случился в тот момент, когда мы дедуцировали каузальную веру из убеждения в тождестве

прошлого и будущего. Долгое время - а и размышляю над этими проблемами не менее 13 лет — мне капалось, что представленные выше схемы исчерпывают ситуацию. Но, столкнувшись с разного рода эпифеноменалистскими парадоксами, я понял, что необходимо перепроверить свои выводы. И в процессе этой перепроверки — а было это летом 2005 г., — к которой меня к тому же подталкивал аргумент «каузальных траекторий» (его мы обсудим позже), я действительно заметил некое упущение, ликвидация которого существенно прояснила ситуацию.

Упущение было связано с тем, что представленный ранее анализ истоков каузальной веры проводился на поверхностном уровне. Мы видели, что вера в то, что каждое событие имеет наблюдаемую в опыте причину, т. е. необходимый коррелят в прошлом, проистекает из веры в тождество прошлого и будущего. Но продумаем еще раз ее специфику: вера в тождество прошлого и будущего означает, что точное повторение прошлого События (состоящего из множества сингулярных событий) сопровождается верой в повторение того События, которое следовало за ним. При этом «прошое событие» не есть темпоральная плоскость, содержащая все его одновременные компоненты. Как уже отмечалось, воспроизводя в уме какое-то событие, мы всегда воспроизводим темпоральный ряд, в который оно встроено, ряд, включающий предшествующие события. Ни одно событие не мыслимо вне такого ряда и, стало быть, включает его в себя.

Это означает, что ожидание некоего события с необходимостью возникает только тогда, когда наличное событие совпадает с неким событием в прошлом (за которым следовало такое же событие, какое теперь ожидается нами) не только относительно его мгновенно-наличного содержания, но и относительно ряда событий, предшествующих этому мгновению. Отсюда следует, что вера в наличие причины у каждого события, возникающая из веры в тождество прошлого и будущего, имеет такую структуру: мы верим, что каждое событие *C* имеет некий событийный коррелят в прошлом, *A* такого рода, что повторение этого *A* в том же самом ряду его собственных antecedentных коррелятов (*B*, *E* и т. д.) всегда будет приводить к событию *C*. Но, как я опять-таки уже показал ранее, ни из чего нельзя вывести, что в другой ситуации у *C* не может оказаться другой причины (вера в тождество прошлого и будущего исключает встроенность в разные событийные ряды лишь совокупных событий). Соответственно, и повторение *A* может произойти с другими его antecedentными коррелятами. И в таком случае мы не обязаны будем верить, что появление *A* приведет к *C*. Иными словами, мы должны считать, что при-

чины событий с необходимостью проявляют свою каузальность лишь при наличии у них тождественной каузальной истории. Таким образом, мы допускаем возможность влияния прошлого на ход событий. Одинаковая событийная конфигурация может вызывать у нас разные ожидания, определяемые различием предшествующих событий. Но чтобы приписывать прошлому реальную действенность, надо допускать, что его содержание и различия не исчезают в настоящем (трудно помыслить, как иначе оно может оказывать влияние). Между тем мы все время предполагали, что в наличном «публичном» пространстве имеется тождество событийных комплексов (какие-то различия в совокупных событиях, конечно, имеются, но по определению они нерелевантны для локальных каузальных рядов). Значит — если оставаться на уровне локальных событийных комплексов, — упомянутая каузальная действенность прошлого может реализоваться лишь в ситуации сохранения его содержания на ненаблюдаемом публично, «приватном» уровне. При отсутствии такого приватного уровня прошлое лишается возможности влиять на настоящее и каузальная нить событий может мыслиться по обычной схеме однозначной связи причины и действия.

Мы, однако, рассуждаем о существах, у которых, как мы знаем на собственном опыте, присутствует приватное измерение. И теперь ясно, что оно действительно может играть роль в детерминации нашего поведения, при том что каждое событие нашей жизни имеет публичную, физическую причину. Все дело в том, что эти физические причины необходимо вынуждают наше поведение лишь при условии их встроенности в те ряды событий, которые имели место в нашем прошлом, а оно оседает в образной памяти, фундирующей объекты наших интенций, т. е. в частных ментальных состояниях. Вот мы и отыскали выход из старого лабиринта ментальной каузальности. Мы сохранили физический детерминизм, так как знания полной физической истории человека достаточно для предсказания его поведения, и вместе с тем признали каузальную роль ментального: физическое прошлое сохраняется только на уровне содержания интенциональных состояний, а поскольку мы признаем возможность влияния этого прошлого на поведение, то мы должны допускать каузальную действенность подобных состояний⁴⁹.

Сразу хочу устранить одно возможное сомнение: здесь можно было бы попробовать возразить, сказав, что все эти рассуждения против эпифеноменализма велись в неверной модальности: я показывал, что для нас естественно верить в реальность ментальной каузальности, а должен был бы показать другое — ее действительную реальность. Но суть

в том, что все наши онтологические воззрения в конечном счете связаны с нашими первичными убеждениями, вроде убеждения в тождестве или соответствии прошлого и будущего, и доказать абсолютную истинность этих убеждений — что позволило бы перевести разговор в объективную онтологическую плоскость — невозможно. Поэтому остается лишь кропотливо описывать наши фундаментальные убеждения и анализировать их следствия. Когда мы делаем это корректно, мы получаем хорошую онтологическую модель. Таким анализом — надеюсь — я до сих пор и занимался.

Суммируем результаты, к которым мы пришли. Я пытался показать, что убеждение в тождестве прошлого и будущего, лежащее в основе нашей когнитивной механики и установок здравого смысла, не только не содействует эпифеноменализму, как могло показаться в какой-то момент, но, напротив, позволяет обосновать возможность реальной ментальной каузальности. Мы тем самым наносим новый удар по эпифеноменализму с позиций здравого смысла, так как возможность ментальной каузальности была показана посредством анализа основоположения здравого смысла. Одновременно мы восстанавливаем в правах еще одно положение здравого смысла (в правомочности которого мы усомнились на предыдущем этапе) — о наличии сознания у других людей, которое, как мы видели, тоже ведет к отрицанию эпифеноменализма.

Проблемность изложенных аргументов, как я уже отмечал, состоит в том, что они могут звучать непривычно даже для искушенного философа. Дело в том, что феноменологический анализ базовых «онтологических установок» (как я назвал их в вводной части книги), а также связей между ними, пример которых они являют, до сих пор оставался плохо освоенным — у меня есть предположения, почему это так, но сейчас я не буду рассуждать на этот счет. К тому же, хотя в современной аналитической литературе очевиден интерес к феноменологическим разработкам Гуссерля и его последователей, боюсь, что интеграция двух традиций еще не настолько велика, чтобы приведенные выше «феноменологические» пассажи были восприняты аналитическими философами, к которым в первую очередь и обращена эта книга. К счастью, однако, эпифеноменистские представления могут быть опровергнуты в том же ключе, но как бы с другой стороны — с помощью вполне аналитического довода.

Довод, о котором сейчас пойдет речь (Б.Б. Ким назвал его «аргументом каузальных траекторий»), пришел мне в голову осенью 2004 г. при размышлении над «Сознающим умом» Д. Чалмерса. Недолго думая, я познакомил с ним Чалмерса,

в результате чего между нами возникла любопытная дискуссия, о которой я еще скажу.

Сам аргумент довольно прост. Сначала предлагается принять посылку о том, что одинаковые события могут иметь разные причины. Она была востребована мной и в «феноменологическом» блоке. Важно подчеркнуть, что речь идет о реальной, а не о логической возможности, т. е. не просто о мыслимости такого положения вещей, а о том, что мы реально можем встретить его примеры в опыте. Эта посылка не так уж невинна — но трудности в ее истолковании можно отвести⁵⁰. К тому же она согласуется со здравым смыслом, допускающим подобную вариативность причин. Скажем, играющие в бильярд не сомневаются, что бильярдный шар может получить одно и то же ускорение вследствие различных движений и соударений; не сомневаются люди и в том, что любая вещь могла бы попасть туда, где она находится, разными путями, и т. п. Скажу еще, что хорошей иллюстрацией работоспособности этой посылки может послужить знаменитый клеточный автомат Life британского математика Дж. Конвея, имитирующий детерминистический универсум⁵¹. В этой симуляции из определенного состояния вытекают строго определенные другие состояния, но само произвольное исходное состояние может быть результатом разнообразных предшествующих событий.

Вторая посылка состоит в констатации того, что моя память отражает события, происходившие со мной в прошлом, т. е. отражает мою каузальную историю. Никаких вопросов она вызывать не должна — отклонения, конечно, всегда возможны, но мы говорим о норме. Столь же несомненно и то, что мои воспоминания - рассматриваемые как приватные качественные состояния — образуют основу моих убеждений, желаний и других осознанных состояний.

Третья посылка указывает на корреляцию моего поведения и моих убеждений и желаний. Эта посылка тоже выглядит самоочевидной. Важно, что ее не отрицает и эпифеноменалист, который, правда, уверен, что эта корреляция является поверхностным явлением, продублированным на более глубоком уровне нейронных связей.

Но сейчас мы покажем, что для такой уверенности нет оснований. Дело в том, что при соединении всех трех посылок мы получаем следующую картину. Если реально, что одно и то же событие или комплекс событий могли бы порождаться разными причинами и если мой мозг в данный момент времени может рассматриваться как громадный комплекс физических событий (некоторые составляющие которого могли бы иметь и иные причины), то очевидно, что он мог оказаться

в данном состоянии в результате самых разных предшествующих событий. Иными словами, мой мозг (и я как физическое существо в целом) мог иметь иную каузальную историю, нежели та, что действительно была у меня и составляла мою жизнь. Но если бы у меня была иная история, то, как следует из второй посылки, я обладал бы сейчас другими воспоминаниями и эти другие воспоминания фундировали бы иные убеждения и желания. Поскольку же мои убеждения и желания, при прочих равных условиях, скоррелированы с моим поведением (третья посылка), то, обладая другими убеждениями и желаниями, я вел бы себя иначе, чем веду себя в данный момент при наличии той истории, которая у меня была.

Итак, мы вывели, что, будучи тем же самым в физическом смысле, я реально мог бы вести себя иначе, чем веду. И различие в моем поведении можно было бы объяснить только наличием других (по сравнению с моим действительным состоянием) убеждений и желаний, взятых с их качественной стороны. Но если это так, то и мое нынешнее поведение должно отчасти объясняться тем, что я наделен такими-то и такими-то ментальными состояниями, и никакой анализ моих физических кондиций не может позволить с точностью предсказать, что именно я буду делать в следующий момент времени.

Все это означает, что качественные аспекты, конституирующие содержания моих убеждений и желаний, играют реальную каузальную роль в моем поведении — и эпифеноменализм ложен.

Когда я изложил Чалмерсу этот аргумент, первая его реакция была такова: при данном физическом состоянии моего мозга я могу, заявил он, обладать лишь одним набором квалиа⁵². Это не что иное, как знаменитый тезис супервентности, с которым согласен и Чалмерс. Но, разумеется, аргумент, предложенный выше⁵³, направлен и на опровержение данного тезиса⁵⁴, точнее на отрицание одной из разновидностей супервентности — локальной супервентности по отношению к качественным состояниям⁵⁵. И все же: нельзя ли предположить, что по мере сближения каузальных траекторий, составляющих мою действительную и мою возможную историю и ведущих к совпадающему моменту моего нынешнего физического состояния — или (если представить тот аргумент в виде мысленного эксперимента, в котором фигурируют двойники с разной историей) по мере сближения меня и моего будущего двойника⁵⁶, — будет происходить постепенное стирание наших ментальных различий? Абстрактно говоря, это, конечно, возможно, но почему, собственно, мы должны это допускать и как это можно мыслить? Ведь когда в конце концов наши ква-

литативные состояния окажутся неотличимы — при разной каузальной истории, — один из нас (т. е. либо я, либо мой двойник) будет корректно представлять свое прошлое, а другой — нет, но априори не видно достаточного основания, почему, к примеру, именно я, а не мой двойник должен быть корректен в своих воспоминаниях. Чалмерс в нашей переписке опробовал несколько иную тактику. Он заявил, что «двойники, скорее, будут иметь что-то вроде амнезии, забывая те аспекты их прошлого, которые отличаются друг от друга»⁵⁷. На это, однако, можно возразить, что в таком случае им придется забыть очень многое или даже почти все, если двойников очень много, а ведь это тоже реально возможно. Иными словами, получится, что сознанием а полной мере обладают только те люди, у которых нет двойников. Чем больше у вас двойников, тем больше вы и они должны быть похожи на зомби, хотя эти двойники никак с вами не связаны. Это странный вывод, который выглядит как доведение до абсурда всей этой линии возражения. И подобный вывод не согласуется с естественными установками нашего сознания, атрибутирующими качественные состояния другим людям вне зависимости от того, имеются ли у них физиологические двойники.

Впрочем, не будем увлекаться разбором возможных возражений и вернемся лучше к тем результатам, которые мы получили. Главный результат: есть все основания считать, что наши приватные качественные состояния оказывают реальное влияние на наше поведение. И это реальное влияние не связано с тем, что данные состояния просто тождественны физическим процессам в мозге и оказывают воздействие именно в силу такого тождества. Мы выяснили, что они не только не тождественны таким состояниям, но и не находятся к ним в отношении супервентности, т.е. не являются чем-то таким, что однозначно соответствует определенным состояниям мозга.

Если этот вывод верен, то он сразу заставляет пересмотреть ряд распространенных представлений о связи мозга и сознания. В частности, данный вывод ставит под сомнение возможность отыскания фиксированных нейронных коррелятов сознательных состояний, особенно если мы говорим не о психических модусах, а об их интенциональных содержаниях, и прежде всего о ментальных образах. Мы также получаем ключ к решению некоторых головоломок, связанных с личным тождеством. К примеру, в книге «Глаз разума», опубликованной в 1981 г. под редакцией Деннета и Хофштадтера, Деннет задается вопросом о том, можно ли утверждать, что в случае копирования молекулярной структуры какого-то человека и воссоздания ее в другом месте (при уничтожении оригинала) будет воссоздан тот же самый — или всего лишь такой же -

человек⁵⁸. При этом Деннет исходит из того, что физическое воспроизведение в любом случае воспроизведет и внутренний мир. Однако мы видели, что это не так. Таким образом, этот мысленный эксперимент беспредметен. Даже если разобрать кого-то на молекулы и через мгновение вновь собрать его из них же, мы не сможем вернуть его уникальное Я.

И в этой связи, т. е. в связи с рассуждениями об уникальном качественном субъекте, нельзя не отметить, что полученные выводы открывают перспективы в решении «трудной проблемы сознания». Читатели уже, возможно, забыли, что книга посвящена именно этому, — но теперь у нас есть повод вернуться к главной теме, и вернуться если не победителями, то по крайней мере с шансами на успех.

«Трудная проблема», напомню, состоит в том, чтобы объяснить, почему функционирование мозга не происходит «в темноте», почему оно вообще сопровождается сознанием, квалиа. Предварительный ответ может состоять в том, что оно сопровождается сознанием просто потому, что без продуцирования сознательных состояний мозг не может функционировать так, как он функционирует, т. е. не может генерировать то поведение, которое он генерирует, так как сознательные состояния участвуют в генерации этого поведения.

Но тут мы сразу сталкиваемся с новой сложностью. Почему мозг не может обходиться собственными, физическими силами? Зачем ему нужно сначала породить квалиа? И кажется, что этот новый вопрос возвращает нас к старой проблеме. И все же это не так. Можем ли мы быть уверены, что мозг в состоянии обходиться собственными силами? Если да, если сугубо физическая система может имитировать человекоподобное поведение, то, в частности, эта задача может быть под силу роботам с надлежащим ИИ. Некоторые авторы, пишущие на эти темы, и прежде всего Сёрл, а в каком-то плане и Чалмерс, уверены, что поведение, подобное человеческому, и правда можно сгенерировать чисто механическими устройствами.

Но так ли это? Нельзя ли попробовать доказать, что никакая чисто физическая машинерия не в состоянии продублировать человеческое поведение, пройти так называемый «расширенный», или «тотальный», тест Тьюринга? Представим, что мы создаем робота-зомби, который должен будет успешно проходить такой тест, т. е. робота, поведенчески неотличимого от обычного человека. Посмотрим, как он будет работать. Спросим его, к примеру, нравится ли ему вкус минеральной воды. Удовольствие, получаемое от минеральной воды, — это качественное состояние. Робот лишен таких состояний. Поэтому он должен эмулировать свое удовольствие, а также

и другие квалиа — иначе он не пройдет указанный тест. Для возможности такой эмуляции необходимо допустить, что каждое качественное состояние заменено у него каким-то физическим состоянием и эти состояния могут фиксироваться в его вербальных отчетах. Итак, если роботы-зомби в указанном смысле возможны, то все квалиа можно заменить физическими аналогами, играющими равнозначные функциональные роли. Вопрос в том, можно ли найти точные физические субституты качественных состояний?

Базовые качественные ментальные состояния интенцио-нальны, и их функционирование зависит от того, каким содержанием, контентом они наполнены. Контент подобных состояний не что иное, как ментальные образы, переносимые в настоящее из прошлого опыта, из памяти. Таким образом, робот, эмулирующий человеческое поведение, должен быть наделен физическими компонентами, позволяющими ему реагировать на наличную ситуацию сообразно его индивидуальной истории. Кажется, что конструирование такого робота не является невозможным. Но есть одна деталь, которая уже не раз отмечалась в этой главе. Все мы признаем, что локальные события — простые или комплексные — могут причиняться разными каузальными рядами. Я не вижу оснований не распространять этот вывод на нашего гипотетического робота. Такой робот мог прийти к своему нынешнему состоянию (которое можно описать как некое комплексное событие) разными путями. Но, независимо от того, как он пришел к нему, он, будучи чисто механической конструкцией, должен в этом состоянии одинаково реагировать на свое окружение.

Итак, для поведения нашего робота-зомби значима не его индивидуальная история, а исключительно его наличное состояние. Конечно, это наличное состояние отражает события, происходившие с ним в прошлом. Но оно с тем же успехом могло бы отражать и совершенно другие прошлые события. Напротив, человеческая память — по крайней мере эпизодическая память, задействованная в рациональном поведении — имеет индивидуализированный характер.

Индивидуальное прошлое не растворяется в нейтральных нейронных кодах, а продолжает существовать на уровне ментальных образов. И именно в таком качестве оно — это прошлое — оказывает влияние на наше поведение. Поскольку эта особенность, как мы видим, не может быть репродуцирована на механическом уровне, проект создания зомбиподобных роботов, похоже, не имеет перспектив.

Итак, если мои рассуждения верны, то ни какой-либо робот-зомби, ни мозг, лишенный качественных ментальных состояний, в принципе не в силах генерировать поведение, ха-

ракторное для человека. И вот теперь мы серьезно продвинулись на пути решения «трудной проблемы», так как мы близки к выводу, что человеческий мозг, продуцирующий поведение, позволяющее самому человеку проходить тест Тьюринга, с необходимостью порождает сознание.

Впрочем, впереди еще много завалов. И первый из них возвращает нас к проблеме каузальной замкнутости физического, принцип которой, как кажется, нарушает сформулированная позиция. В самом деле, если сознание реально влияет на поведение, то оно реально влияет и на функционирование мозга. Иными словами, мозг как локальная физическая система оказывается недоопределен своим наличным состоянием: для того чтобы понять, в каком направлении оно будет изменяться, надо принимать во внимание и приватные состояния, имеющиеся у меня в данный момент времени. Поэтому физические события, характеризующие состояние моего мозга в последующий момент времени, частично детерминируются нефизическими, качественными причинами, что, собственно, и означает отрицание принципа каузальной замкнутости физического.

Все это могло бы показаться не таким уж страшным, особенно если учесть, что некоторые интерпретации квантовой механики, рисуящей нам картину мира на самом фундаментальном уровне, по крайней мере допускают участие сознания в детерминации физических процессов. Но дело в том, что принцип каузальной замкнутости физического отчасти генерируется базовым убеждением в тождестве прошлого и будущего — и так просто от него не избавиться, во всяком случае когда мы начинаем рассуждать о других людях и о других сознаниях, пытаюсь определить каузальную роль последних.

Но не противоречит ли то, что я только что сказал по поводу принципа каузальной замкнутости, проделанному выше анализу следствий нашего убеждения в тождестве прошлого и будущего? Ведь мы выяснили, что, хотя относительно каждого физического события можно сказать, что оно имеет физическую причину, сама эта физическая причина должна пониматься как некий ряд событий, уходящий в прошлое, а не как изолированное наличное событие. Если же брать изолированное событие, то его каузальный потенциал может оказываться недостаточным для классификации его в качестве причины (и, соответственно, требовать каких-то дополнений со стороны квалиа). И если понимать принцип каузальной замкнутости как положение о том, что каждое физическое событие имеет физическую причину в непосредственном прошлом, то может возникнуть ощущение, что в такой формулировке этот принцип попросту неверен. Но тогда

нечего опасаться нестыковки принципа каузальной замкнутости и сделанных выводов об ошибочности эпифеноменализма. Между тем, выше я утверждал, что основания для опасений есть.

Сейчас мы увидим, что такие основания действительно имеются, и что принцип каузальной замкнутости верен именно в той формулировке, которая могла показаться нейтрализованной. Развернем еще раз весь сценарий предложенной аргументации. Допустим, что я мог прийти к своему наличному физическому состоянию другим путем, и если бы это произошло, я повел бы себя иначе, чем веду сейчас. Важно при этом, что наличное физическое состояние мира, в котором я повел бы себя иначе, будучи физически тождествен себе, какой я есть в настоящий момент времени, не могло бы в этом случае быть — точнее, мы не можем в это поверить — совершенно тождественным тому состоянию, в котором мир находится сейчас.

И вот почему. Подумаем сначала, могу ли я верить, что в настоящий момент времени в таком мире, в котором я существую, люди могли бы вести себя иначе, чем они будут вести себя в следующее мгновение? Нет, потому что они могли бы вести себя иначе лишь в случае иной каузальной истории, но иная каузальная история соответствовала бы другим состояниям мира в предыдущие моменты времени. Иначе говоря, если бы я допускал, что люди, существующие в каком-то мире, реально могут в следующий момент времени вести себя по-разному именно в таком мире, то я верил бы, что подобный мир в целом мог прийти к своему наличному состоянию другим путем — но это противоречит принципу тождества прошлого и будущего, который, разумеется, имеет силу в обоих направлениях.

Таким образом, мир, в котором физические двойники людей будут вести себя иначе, должен отличаться от мира, в котором они реально существуют. Причем я должен мыслить эти отличия так, чтобы они могли быть даны в моем опыте. Но я не могу допускать, что отличия, о которых идет речь, не выходят за пределы моих частных состояний, — ведь они должны быть даны в опыте и других людей. Поэтому я должен считать, что отличия, скоррелированные с разным поведением физически одинаковых людей, сами имеют физический характер. Этот вывод я могу перенести и на себя. Наличный мир (т. е. совокупное событие, которым он, собственно, и является) и мир, в котором я был бы таким же в физическом плане, но вел себя иначе, должны отличаться в каких-то других физических отношениях. Поскольку локально они тождественны — так как в обоих есть один и тот же я (рас-

смаатриваемый исключительно со своей физической стороны) в одном и том же окружении, — то это отличие должно проявляться за пределами его совпадающих мест. Так вот, разницу в поведении меня и моего гипотетического двойника можно в принципе объяснить ссылкой на эти удаленные физические различия.

В общем, можно сказать, что мое будущее поведение все же имеет физические корреляты в наличном состоянии мира, разве что эти корреляты не локальны, а разнесены со мной в пространстве. Тем не менее этого вполне достаточно, чтобы спасти принцип каузальной замкнутости физического в его традиционном понимании. Но тогда мы возвращаемся к старой проблеме, хотя в несколько иной формулировке: подтверждая глобальную супервентность ментального на физическом, не приведет ли тот тезис к восстановлению эпифеноменализма? Если мое поведение все же имеет физические корреляты, пусть и нелокальные, то какую роль играют квалиа?

На этот вопрос можно предварительно ответить следующим образом. Во-первых, сами квалиа очевидным образом скоррелированы с теми физическими коррелятами моего поведения. И они тоже нелокальны, правда в несколько ином смысле: они вообще не даны в публичном пространстве, а не только в данном локусе⁵⁹. Во-вторых, квалиа все же связаны с локусом мозга, так как зависят от мозга. В этом смысле они, наоборот, локальны. В-третьих, нелокальные физические корреляты моего поведения прямо не влияют на мозг в момент, предшествующий переходу мозга из одного состояния в другое, а значит, и вообще напрямую не влияют на него — если бы они влияли, нельзя было бы говорить о моем физическом тождестве с двойником, так как его поведение сопряжено с другими нелокальными коррелятами и их влияние было бы другим, что неизбежно приводило бы к обнаружению нашей физической неодинаковости.

Объединяя эти моменты, получаем вот что: поскольку нелокальные физические корреляты моего поведения не оказывают прямого влияния на физическую систему мозга, тогда как качественные состояния, в силу их сопряженности с мозгом, могут делать это, то глубинными причинами нашего поведения, поскольку оно вообще зависит от нелокальных, т. е., в частности, от качественных факторов, сознательных состояний, являются именно последние, а не нелокальные физические корреляты. Если бы произошло чудесное вмешательство в ход физических процессов и за мгновение до моего поступка, скоррелированного с нелокальными физическими факторами, эти корреляты были бы уничтожены, то я повел бы себя точно так же, как если бы они остались на месте —

так как настоящими движущими причинами моего поведения, помимо локальных физических факторов, были вовсе не эти нелокальные корреляты, а качественные, приватные ментальные состояния. Иначе говоря, мы видим, что в этой ситуации на роль эпифеноменов моего поведения претендуют, скорее, не сознательные состояния, а физические факторы.

Но здесь, конечно, надо быть осторожным в формулировках. Нелокальные физические корреляты моего поведения все же могут быть названы его физическими причинами, так как их статус соответствует базовым характеристикам физических причин — «чудесное вмешательство», о котором шла речь выше, нереально⁶⁰. И принцип каузальной замкнутости, который мы защищаем, требует именно признания наличия у каждого физического действия физической причины. С другой стороны, эти корреляты все же не вполне обычные причины, так как они работают через качественные состояния, которые тем самым оказываются условиями их действенности, а значит, чем-то более фундаментальным — что и позволяет трактовать первые как некое подобие эпифеноменов, хотя и не более чем подобие.

Одним словом, если мы рассматриваем мозг как локальную систему, то у нас есть все основания говорить, что ее поведение определяется в том числе и качественными ментальными состояниями — локальный эпифеноменализм ложен. Если же мы берем мир в целом, то мы могли бы говорить об эиифеноменальности подобных состояний — хотя с равным основанием можно было бы считать каузально неэффективными, наоборот, нелокальные физические корреляты нашего поведения. Такая двойственность, правда, не есть достоинство. И в действительности оба упомянутых варианта неверны. Так что если мы хотим онтологической точности, то можно сказать, что ментальные состояния являются условиями действенности нелокальных физических причин — большего здесь и не требуется⁶¹.

И все же без дальнейших уточнений не обойтись. Хотя с онтологической точки зрения обозначенная позиция достаточно ясна и хотя я уже привел некоторые аргументы в ее пользу, они, по правде говоря, не являются решающими. А если они не решающие, то мы оказываемся в ситуации, где могут воспроизвестись все старые вопросы — и «трудная проблема» вновь ускользнет от своего решения.

Все дело в том, что пока я так и не привел убедительного доказательства необходимости сознания, приватных состояний, для реализации нелокального физического влияния. Если такое доказательство не будет найдено, то наличие сознания по-прежнему будет казаться чем-то необъяснимым

и по-прежнему будет неясно, почему наши действия не могут детерминироваться чисто физически.

Как и раньше в подобных запутанных случаях, обратимся к феноменологическому анализу наших базовых убеждений. Ранее мы видели, что (1) у нас есть все основания считать, что определенные сингулярные события каузально связаны с другими сингулярными событиями и их рядами, образующими автономные цепи. Иначе говоря, мы исходим из того, что каузальные ряды, обнаруживающиеся в нашем мире, локальны и поэтому реализуемы в других возможных мирах (хотя, конечно, наверняка лишь в тех, которые близки нашему настолько, что их состояния можно приравнять к будущим состояниям нашего мира). Мы также выяснили, что (2) присущий нам механизм ожидания событий устроен так, что на эти ожидания влияет имеющееся у нас представление о том, какой именно каузальный ряд привел к той причине, действия которой мы в данный момент ожидаем. Кроме того, было показано, что (3) различия в прошлых каузальных рядах такого рода неизбежно экстраполируются нами на наличное состояние мира и что эти наличные различия имеют нелокальный характер по отношению к рассматриваемым нами рядам.

И вот тут решающий момент. Нам надо понять, почему при такой структуре ожидания мы не считаем, что все физические события не полностью детерминированы локальными факторами. Ответ состоит именно в том, что — обобщая опыт реальных каузальных зависимостей — мы изначально исходили из того, что в нашем мире физические события детерминируются локальными причинами, а принятие вышеописанной схемы отменяло бы это допущение, обозначенное выше как пункт (1). Поэтому, конструируя свои ожидания относительно поведения тех или иных материальных систем, обычно мы игнорируем вопрос, каким образом они пришли к нынешнему состоянию, полагая, что физическая причинность действует локально и на основе наличных физических обстоятельств. Вместе с тем мы не можем исключать, что если бы прошлый опыт аккумулировался внутри самих каузальных рядов на приватном уровне, этот опыт мог бы оказывать влияние на ход физических событий. Так оно и выходит. И лишь в этом случае мы можем соотносить происходящие события с нелокальными физическими коррелятами приватных ментальных состояний — а фактически соотносить их с состоянием мира в целом. Но нарушения естественно принимаемого допущения автономности каузального ряда, связанного с историей, к примеру моего организма (возникающего в конечном счете от пересечения чисто физических рядов, автономность которых не вызывает сомнений), все равно не происходит, ибо

нелокальные факторы могут рассматриваться как определяющие причины моего поведения лишь при допущении, что они действуют через приватные ментальные состояния, отражающие историю развертывания этого каузального ряда.

При понимании того, что нелокальные физические причины могут действовать только через локальные в одном смысле — т. е. сопряженные с данным автономным рядом — и нелокальные в другом смысле — приватные — факторы, становится очевидной невозможность обойтись без квалиа, без приватного измерения⁶². При этом само приватное измерение, разумеется, может обнаружиться лишь при определенных структурных предпосылках, при определенной организации материальной системы, в которой оно существует.

А чтобы понять, о какого рода организационных принципах здесь может идти речь, отметим, что, поскольку квалиа, как мы выяснили, должны быть связаны с нелокальными физическими явлениями, а не локальность — феномен, хорошо известный квантовой физике, то мы можем попробовать соотнести сознание с квантовыми феноменами (хотя при этом надо стараться избегать разного рода спекуляций). Самый наглядный способ сделать это — вспомнить, что, согласно нашим выводам, события, происходящие в мозге как физической системе, не полностью детерминированы локальными физическими обстоятельствами. И если — гипотетически — в какой-то момент лишить мозг квалиа, то его состояние в следующий момент времени можно мыслить наподобие суперпозиции множества его возможных изменений. Соответственно, роль квалиа в таком случае сопоставима с фактором, вызывающим коллапс волновой функции, — но ведь именно эта роль может отводиться качественным состояниям (сознанию) в ряде интерпретаций квантовой механики⁶³. И тут я должен напомнить, что со времен статей Пенроуза и Хамероффа — хотя, конечно, у них в этом плане было немало предшественников — тема связи сознания с квантовыми процессами занимает умы авторов самой разной специализации. Впрочем, исследование возможных квантовых основ сознания находится в самой начальной стадии, хотя некоторые из выдвинутых гипотез уже пошатнулись⁶⁴. Поэтому я не буду углубляться в частности⁶⁵. Рискну лишь предположить, что, несмотря на то, что объективно мозг в силу его температурных параметров является неблагоприятной средой для квантовых макроявлений, поиски связи сознания и квантовых феноменов — перспективное направление. Что-то здесь обязательно будет найдено.

С философской точки зрения интереснее то, что квантовые макроэффекты в мозге, от которых, вероятно, зависит воз-

никновение сознания, в силу специфики их протекания должны изначально обнаруживаться на клеточном или — если (что весьма вероятно в силу интегрированности сознательных феноменов) они реализуются в синапсах нейронов — на межклеточном уровне и едва ли могут генерироваться большими системами⁶⁶. Это важно, потому что если это действительно так, то квалиа или протоквалиа могут порождаться уже, к примеру, на уровне небольшой коалиции нейронов. А это, в свою очередь, позволяет предположить, что качественные состояния могут быть присущи даже примитивным существам, если органические структуры, образующие их клетки, обладают характеристиками, которые порождают квалиа в нейронных структурах мозга.

Не менее принципиальным представляется и другое следствие низкоуровневой природы квалиа: квалиа, существующие, возможно, даже при компактных коалициях нейронов, не обязательно должны быть связаны с сознанием в его обыденном понимании. Иными словами, есть серьезные основания прислушаться к уже известным нам идеям Н. Блока, предлагающего различить два типа сознания: феноменальное сознание и сознание доступа. Феноменальное сознание имеется там, где имеются квалиа. А сознание доступа обнаруживает себя в тех случаях, когда те или иные квалиа оказываются доступными различным системам мозга или когда они могут оказывать влияние на функционирование этих систем. Чем больше такое влияние, тем интенсивнее осознание соответствующих квалиа.

На первый взгляд, правда, не совсем ясно, как соотносится понятие неосознанных — в смысле доступа — квалиа и бессознательных состояний, взятых в классическом фрейдовском смысле. Ведь, к примеру, бессознательные желания, похоже, могут оказывать влияние на поведение, и тем самым они скорее подпадают под понятие сознания доступа⁶⁷. С другой стороны, очевидно, что такие состояния должны соответствовать именно феноменальному сознанию без доступа. Решение, возможно, в том, чтобы трактовать сознание доступа как нечто градуированное. Т. е. все квалиа оказывают какое-то влияние (хотя бы потому, что все нейроны, где они, вероятно, порождаются, связаны с другими нейронами). В этом смысле сознание доступа, пусть и слабое, распространяется на все качественные состояния. Но при увеличении их влияния, т. е. при усилении соответствующих нейронных паттернов в глобальном рабочем пространстве (если допустить существование такой амплифицирующей среды), сознание доступа становится таким интенсивным, что эти квалиа выступают как нечто обособленное, отличное от всего остального.

Так вот, фрейдовские бессознательные состояния можно рассматривать как квалиа, достаточно влиятельные для того, чтобы участвовать в формировании поведенческих реакций, но недостаточно влиятельные для того, чтобы представлять как нечто обособленное, т. е. чтобы быть отличенными от других квалиа. Тут, кстати, уместно вспомнить, что в новоевропейской философии под сознанием нередко понимали именно способность различения представлений⁶⁸. Если одно представление отличено от другого, то можно сказать, что оно осознается. Соответственно, блоковскую схему можно модифицировать и превратить в дихотомию феноменальное сознание / сознание-различение (понятие сознания доступа оказывается неоптимальным, так как оно, как видно, должно рассматриваться как нечто континуальное). Далекое не все квалиа отличаются мной от других. Но тут новый вопрос — а кто такой Я? Нужно ли примысливать к этой схеме особого субъекта?

Видимо, да, но в такой процедуре нет ничего нелегитимного. Субъект можно мыслить как систему мониторинга, сопряженную, вероятно, с лингвистическим модулем в так называемом доминантном полушарии⁶⁹ и работающую по триггерно-му принципу: по достижении тем или иным качественным состоянием и коалицией нейронов, его физической базой, определенного влияния в мозге триггер срабатывает и соответствующие квалиа становятся контентом других ментальных состояний, связанных с этой системой. Именно тогда мы можем утверждать, что квалиа различены и осознаны.

Данная схема, являющаяся разновидностью так называемых «высокопорядковых теорий сознания», один из самых известных вариантов которой разработан Д. Розенталем⁷⁰, позволяет разрешить множество старых проблем, связанных со статусом Я. В частности, Я должно быть лишено титула субстанции, так как соответствующие этой инстанции качественные состояния — которые, кстати, участвуют в формировании «онтологических установок», подробно обсуждавшихся нами в этой главе, — так же зависимы от мозга, как и все другие квалиа. Кроме того, очевидно, что Я, истолкованное в указанном смысле, не может быть центром ментальной каузальности — ведь квалиа действительны вовсе не из-за их осознанности.

Не может это Я быть и источником свободных решений. Тут, кстати, надо сделать небольшую остановку. Дело в том, что тема ментальной каузальности часто смешивается с проблемой свободы воли. Но это совершенно разные вещи. Вспомним, что, согласно ранее полученным выводам, реальность ментальной каузальности не противоречит тезису о детерминистичности мира, в том числе мира, взятого исключительно

В физическом аспекте. Проблема же свободы воли — если она не привязывается к тривиальной и встроенной в обыденные контексты здравого смысла констатации, что некоторые наши решения принимаются после обдумывания разных вариантов и выбора лучшего из них (и в этом смысле свобода, конечно, реальна и совместима с детерминизмом, причем дело здесь действительно не обходится без участия Я, так как те варианты должны быть различены), — подразумевает ее индетерминистическое понимание. Лишь при либертарианской трактовке эта проблема по-настоящему интересна. Истоком представления о возможности такого рода свободы служит то, что можно назвать «чувством свободы». И говорит оно о том, что, скажем, через секунду я реально могу поступить так или иначе. Но если мы проанализируем его, мы убедимся в его иллюзорности. Итак: когда возникает такое чувство, мы считаем реально возможным совершение разных действий, к примеру, в ближайшие секунды. Уверенность, что то или иное событие реально возможно, основана на перенесении прошлого опыта на будущее — и в данном случае это именно так, поскольку, думая о возможных действиях, мы перебираем далеко не все варианты, в частности, я не считаю возможным через секунду оказаться в Париже. Но перенесение прошлого на будущее, как мы знаем, предполагает веру в детерминированность событий. Это значит, что в любой конкретной ситуации мы должны допускать только одно реально возможное развитие событий. Ощущение же вероятности, а значит, и реальной возможности разных вариантов берется от ограниченности нашего знания деталей. Таким образом, чувство свободы иллюзорно и возникает лишь при предположении веры в детерминистическую картину мира⁷¹.

Но вернемся к «трудной проблеме». Теперь мы, в частности, уже подготовлены для того, чтобы дать более или менее удовлетворительный ответ на вопрос, почему функционирование мозга сопровождается сознанием, если брать слово «сознание» в узком смысле как синоним сознания-различения. И ответ этот таков: потому что мозг (1) порождает квалиа и (2) имеет в качестве одного из своих блоков систему самомониторинга, продуцирующую качественные ментальные состояния, контентом которых оказываются квалиа, достаточно влиятельные в глобальном рабочем пространстве, чтобы срабатывал триггерный механизм, открывающий им доступ в эту систему.

Эта часть «трудной проблемы», таким образом, допускает вполне прозрачное решение. Ее можно назвать «простой» частью. Но главная ее — собственно «трудная» - часть состоит в вопросе: почему мозг вообще порождает квалиа? Однако

и здесь мы близки к ответу. Мы уже знаем, что мозг порождает квалиа потому, что без квалиа, играющих реальную каузальную роль, мозг просто не может функционировать так, как он функционирует. Но чтобы этот ответ звучал убедительно, надо пояснить, как именно он не смог бы функционировать без качественных состояний. В чем специфика функционирования мозга, наделенного квалиа? Вспомнив недавние рассуждения о невозможности механического дублирования поведения человека, мы поймем, что в общем знаем ответ и на этот вопрос: мозг, наделенный квалиа, может генерировать гораздо более разнообразное поведение, нежели чисто механическая система.

Более разнообразное поведение лучше способствует выживанию⁷². Отсюда и объяснение причин эволюции сознания - его появление уже не выглядит чем-то случайным или таинственным. Но, вопросы, конечно, еще остаются. Я все время повторяю, что мозг порождает сознание, порождает квалиа и т. п. Оспаривать это, учитывая зависимость ментальных состояний от состояний мозга, едва ли возможно. Но, с другой стороны, не получается ли здесь какого-то непонятного круга: мозг порождает квалиа, а они, в свою очередь, влияют на функционирование мозга. Зачем такие сложности?

Ответ может быть таким: хотя мозг действительно порождает квалиа, но каузальная роль качественных состояний шире того нейронного базиса, на основе которого они возникают. Ведь эта каузальная или функциональная роль отчасти определяется окружением организма, взаимодействием с которым наполняет наши приватные формы ментальным контентом. Т. е. возможно, к примеру, что существует жесткая корреляция между данным физическим состоянием мозга и наличием у меня желания вообще. Но, как следует из тех тезисов, которые рассматривались в данной главе, физическое состояние моего мозга не может в полной мере детерминировать, каким именно желанием я обладаю, что именно я желаю. Предмет моего желания зависит от моего окружения и истории взаимодействия меня как физического существа с миром⁷³. Но ведь именно этот предмет в конечном счете определяет мое поведение. Таким образом, порождая сознание, мозг изготавливает для себя полезнейший инструмент удержания индивидуализированного прошлого опыта.

Поскольку порождение сознания мозгом нельзя рассматривать по модели каузальной зависимости, то его следует мыслить, скорее, как создание им приватного пространства для ментальных состояний. Как уже отмечалось, есть определенные основания допускать, что подобное пространство может возникать не только в мозге, но и в гораздо более прими-

тивных системах. И возникает оно при появлении в функционировании таких материальных систем того, что можно назвать квантовыми эффектами. Разумеется, квантовые эффекты существуют во всех системах, но принято считать, что обычно они нейтрализуются на макроуровне. Не исключено, однако, что некоторые макросистемы, если рассматривать их исключительно с физической стороны, расположены вести себя неклассически — на подтверждение этого тезиса направлены сейчас усилия немалого количества исследователей⁷⁴. Именно в таких системах и должно обнаруживаться то самое приватное измерение.

Итак, «трудная проблема» получает следующее решение: квалиа сопровождают функционирование человеческого мозга (а, вероятно, и других неклассических систем), потому что они являются необходимыми условиями продуцирования мозгом такого поведения организма, которое он фактически продуцирует и которое учитывает его индивидуализированный опыт⁷⁵. Поскольку оно является выгодным в адаптационном смысле, неудивительно, что подобные системы были отобраны в ходе эволюции.

Но это еще не все решение. Напомню, что «трудная проблема» включает в себя два вопроса. Первый — почему мозг порождает квалиа, или сознание? Второй — как мозг порождает квалиа? И кажется, что пока я ответил лишь на первый вопрос. В действительности ответ на второй вопрос тоже уже неявно дан, но дело в том, что он не может быть исчерпывающим. Ведь ответить на него — значит определить механизмы порождения мозгом квалиа. И с философской точки зрения здесь можно говорить лишь о том, что приватное измерение, в котором располагаются квалиа, конституируется структурами, благодаря которым материальная система оказывается неклассической, т. е. не полностью определенной локальными физическими факторами, — и может, как мы видели, вследствие этого производить такие действия, которые обусловлены историей этой системы.

Но хотя такое решение не является чистой абстракцией и позволяет отсеять ряд гипотез, таких, как панпсихизм, противоречащий, кстати, здравому смыслу, очевидно, что оно не может быть полным и оставляет непроясненным множество эмпирических деталей. Впрочем, будучи следствием концептуального анализа, оно и не может претендовать на уточнение всех конкретных обстоятельств. Так что оно не конфликтует с программой экспериментальных исследований нейронных основ сознания, а, скорее, подталкивает к ним. И разумеется, это хороший признак, указывающий на реальную перспективность взаимодействия эмпирической науки и философии.

Но вернемся, собственно, к философии. Теперь мне предстоит найти имя для предложенной позиции и пояснить ее отношение к другим, конкурирующим теориям сознания. Наиболее точным было бы, пожалуй, назвать ее «локальным интеракционизмом». Другое имя уже звучало в одной из моих статей — «эмерджентный интеракционизм»⁷⁶. Думаю, оно не нуждается в особых комментариях. Доказывая ложность локального эпи-феноменализма и каузальную действенность сознания, я тем самым сближаюсь с интеракционистской позицией, хотя мы видели, что этот интеракционизм позволяет избежать нарушения принципа каузальной замкнутости физического, и, на мой взгляд, трудно спорить, что это является его несомненным достоинством. Что же касается эпитета «эмерджентный», то он служит тому, чтобы отграничить эту позицию от классического дуалистического интеракционизма, признающего ментальные данности состояниями особой духовной субстанции. Я уже пытался показать, что для этого допущения нет никаких оснований. Самое понятие духовной субстанции возникает из неверной интерпретации феномена Я.

Субстанциализация Я не только неправомерна сама по себе, но и приводит к множеству, мягко говоря, сомнительных следствий, касающихся независимости ментальных состояний от мозга, возможности их сохранения после разрушения организма, и тождества Я.

Особенно любопытны в этой связи положения о тождестве и единстве Я, являющиеся неизбежным результатом трактовки Я как особой субстанции. И любопытны они потому, что могут быть напрямую фальсифицированы опытом. Особенно очевидно это в случае тезиса о единстве Я (если считать это Я носителем всех своих ментальных состояний), разрушаемого данными об раздвоении психики при рассечении *corpus callosum* и других путей, связывающих два полушария⁷⁷, и другими не столь драматичными экспериментами.

Предложенное же выше — и здесь я иду за многими другими — истолкование статуса Я и качественных состояний позволяет легко объяснить все эти экспериментальные данные. И тождество, и даже единство нашей ментальной жизни — условные вещи, существующие до тех пор, пока существует единство физического носителя квалиа. А существует оно до тех пор, пока не прекращается локальное взаимодействие материальных частей, образующих нормальный человеческий мозг. Таким образом, слово «эмерджентный» позволяет подчеркнуть зависимость сознания от мозга, избегая вместе с тем физикалистского искушения трактовать квалиа как нечто супервентное или даже тождественное тем или иным физическим параметрам мозга.

Теперь можно попробовать определить отношение представленной в этой главе концепции к другим теориям сознания. И начать логично с тех, которые подробно рассматривались ранее. В отличие от Деннета, Армстронга и других эли-минативистов, функционалистов и теоретиков тождества я полагаю, что физические функциональные схемы в мозге не исчерпывают качественные состояния, а включают их в качестве внешних по отношению к ним (но внутренних по своей природе, частных) компонентов. Соответственно, я считаю неверным тезис Чалмерса о параллелизме сознательных состояний и физических функциональных схем. Неточным вследствие этого представляется мне и вывод Чалмерса о том, что механическое дублирование функциональных схем живых существ должно приводить к возникновению сознания. Но неточен этот тезис потому, что механически дублировать эти функциональные схемы, т. е. дублировать их исключительно на локальном физическом уровне, попросту невозможно. Вместе с тем, если бы все-таки удалось вычислить алгоритм человеческого поведения (это обязательно должен быть какой-то нестандартный, открытый алгоритм) и инсталлировать соответствующую программу в робота, то этот робот стал бы сознательным существом. Поэтому я не могу согласиться с Аргументом китайской комнаты Сёрла — при внимательном анализе его Китайская комната доказывает прямо противоположное тому, что хотел Сёрл, а именно: она доказывает безальтернативность «сильного ИИ», т. е. тезиса, что инсталляция программы, точно эмулирующей человеческое поведение и позволяющей системе с такой программой проходить тотальный тест Тьюринга, с необходимостью в силу этого приводит к появлению в соответствующей системе качественных состояний. Ошибается Сёрл и в том, что качественные состояния не содержат в себе каузальных возможностей, которые выходили бы за пределы каузальных возможностей порождающего их мозга. Не вполне оправдан и его тезис о каузальном порождении мозгом квалиа.

Таким образом, в системах Сёрла, Деннета и Чалмерса нет, пожалуй, ни одного принципиального положения, касающегося «трудной проблемы», под которым я мог бы подписаться. Можно было бы ожидать, что большие пересечения обнаружатся в системах современных авторов, отстаивающих ин-теракционистские взгляды. Проблема, однако, в том, что системы такого рода не всегда выглядят системными. Так, самый впечатляющий по объему интеракционистский труд последних десятилетий — «Самость и ее мозг» (1977) Поппера и Экклза — содержит множество любопытных наблюдений, но неконкретен по ключевым вопросам. Авторы доказывают,

что без допущения влияющего на мозг «самосознательного ума», или «чистого Я», невозможно объяснить интеграцию чувственных донных и феномен свободы воли. Поппер также затрагивает вопрос о возможном нарушении первого и второго закона термодинамики в результате подобного влияния, но ограничивается самыми общими замечаниями о том, что это нарушение может быть настолько незначительным, что его нельзя будет зафиксировать наблюдениями⁷⁸. Замечу, что принципиальные трудности с этими законами возникают лишь при классической трактовке интеракцио-низма и лишены остроты при том его понимании, которое было предложено в этой главе: напомню, что локальный эмерджентный интеракционизм не отрицает каузальную замкнутость физического на глобальном уровне и, соответственно, может и вообще не конфликтовать с основными постулатами физики.

Рассмотренная в данной книге версия интеракционизма позволяет избежать и так называемой трудности «установления пар» (pairing problem), о которой писал Джеквон Ким⁷⁹. Он доказывал, что классический картезианский дуализм не в состоянии объяснить, почему то или иное физическое событие вызывается определенной духовной субстанцией. Допустим, что на основе контрфактуальных рассуждений мы делаем предположение, что некое желание, возникающее у подобной субстанции, т. е. души, приводит к физическому событию А, но не к одновременному физическому событию В. Но чем отличается отношение этой души к событию А от ее отношения к событию В? В аналогичном случае физической причинности на такой вопрос всегда можно ответить, указав на определенные пространственные связи, имеющиеся между причиной и действием. Но духовная субстанция, по дефиниции, лишена пространственности. Поэтому каузальные отношения между ней и конкретными физическими событиями оставались бы, полагает Ким, чем-то необъяснимым.

Однако, даже если этот аргумент верен в отношении картезианского интеракционизма, допускающего принципиальную независимость души от физического⁸⁰, он не работает против эмерджентного интеракционизма, согласно которому ментальные состояния, онтологически отличные от физических процессов в порождающей их материальной системе, могут быть каузально действенными по отношению к ней из-за того, что они размещаются в ее приватном измерении и отражают ее физическое прошлое⁸¹. Мое желание поднять руку приводит к тому, что перемещается не рука, скажем, какого-нибудь художника-экспрессиониста, а моя рука, потому, что

данное желание встроено в приватное пространство и историю физической системы, частью которой является моя рука, и ни в какую другую.

Но вернемся к интеракционистским теориям. Наиболее систематичным из новых интеракционистов является, по моим наблюдениям, Уильям Хескер, автор уже упоминавшейся книги «Эмерджентная самость». В своей работе он, кстати, показывает, что взгляды некоторых других формально близких ему авторов, и прежде всего Р. Сперри, либо вообще некогерентны из-за радикальных противоречий, либо выражены так, что точный смысл сказанного остается неясным⁸². Однако и теории самого Хескера не отличаются большой определенностью⁸³, особенно когда речь у него заходит о природе ментальной каузальности, — а те детали, которые все же можно обнаружить в его тексте, представляются мне спорными. Хескер считает необходимым заключить, что наш мозг порождает не просто сознательные состояния, а единое поле сознания, или субстанциального индивида, причем этот индивид обладает свободой воли в либертарианском смысле, несовместимом с детерминизмом. Мы уже видели, что понятие о такой свободе воли иллюзорно, так же иллюзорно, как представление о безусловном единстве субъекта.

Под занавес главы, в которой была предпринята попытка сформулировать решение «трудной проблемы сознания», логично оценить идеи философа, заявившего о существовании неустранимого «провала» в объяснении отношения сознания и мозга — Дж. Левина. Всякое объяснение, полагает Левин, включает некую дедуцию, но никоим образом из характеристик физического состояния человеческого мозга нельзя дедуцировать, что его функционирование будет сопровождаться такими-то и такими-то квалиа. Это и есть «провал в объяснении». Спорить с тем, что квалиа логически невыводимы из физических характеристик мозга, невозможно, но, вообще говоря, любое действие невыводимо из представления о его причине. А здесь все усугубляется тем, что между мозгом и квалиа нет однозначного каузального отношения. Кроме того, в силу приватности, качественные состояния других существ не могут быть даны нам. Эти обстоятельства и создают ощущение «провала в объяснении». Между тем Левин, похоже, смешивает тезис о невозможности выведения качественных характеристик ментальных состояний с положением о невозможности объяснения существования квалиа вообще⁸⁴. Мы видели, однако, что для отождествления этих тезисов нет оснований, — и объяснить, почему мозг порождает квалиа, вполне реально. В этом центральном вопросе у нас нет провала в объяснении.

Соответственно, нет причин соглашаться и с пессимистическими утверждениями Колина Макгинна, рассуждающего о нашей «когнитивной замкнутости» в понимании отношения между сознанием и мозгом⁸⁵. Вся эта книга была нацелена на доказательство того, что имеющиеся в этих вещах трудности могут быть преодолены.

Не могу при этом не отметить, что если решение «трудной проблемы», предложенное выше, может претендовать на истинность, то оно подталкивает к пересмотру оценки так называемых «легких проблем сознания», о которых я говорил в первой главе и в главе о Чалмерсе. Ведь само понятие «легкие проблемы» было введено Чалмерсом потому, что, по его мнению, мы хорошо представляем, в каком направлении их можно решать. Речь при этом шла о таких проблемах, как отыскание нейронных коррелятов сознания, построении функциональных и компьютерных моделей различных когнитивных операций и т. п. Но дело в том, что ясный алгоритм решения подобных проблем мог бы иметься у нас, лишь если бы наши приватные состояния не оказывали влияния на функционирование мозга и, соответственно, на поведение человека⁸⁶. Тогда, исследуя человеческое поведение и его порождающие механизмы в мозге, равно как и нейронные основы сознательных состояний, можно было ограничиваться стандартными научными методами от третьего лица.

Но раз сознание влияет на поведение, эти методы оказываются недостаточными. Некоторая пробуксовка когнитивной науки, и в частности разработок в области искусственного интеллекта, может объясняться тем, что эти проекты были скроены по образцу решения «легких проблем» и при предположении возможности чисто механической реализации соответствующих когнитивистских и иных моделей. Это не значит, конечно, что такие проекты вообще бесперспективны — я не сомневаюсь, что в этих областях нас ждут удивительные успехи. Но я также убежден, что они произойдут на фоне серьезных и, скажем так, идеологических трансформаций. Примером осознания необходимости такого рода изменений в сфере исследований по ИИ может служить недавняя книга Дж. Хоукинса⁸⁷. Раздаются в последнее время призывы к капитальной методологической перестройке и в когнитивной нейронауке⁸⁸. Так или иначе, но «легкие» проблемы могут оказаться — и даже наверняка окажутся — труднее «трудной». Но это еще не повод для расстройства.

Примечания

¹ К тому же это решение было предложено еще до появления термина «трудная проблема сознания».

² О понятии каузальной релевантности — см.: Черняк А.З. К вопросу о каузальной эффективности ментальных свойств // *Философия сознания: классика и современность*. М., 2007. С. 105-113.

³ Кроме того, теория тождества, как доказывал, в частности, Дж. Ким, не избавляет от опасности эпифеноменализма (ведь ментальные состояния все равно могут лишиться действительности в качестве таковых), хотя некоторые традиционные возражения против эпифеноменализма теряют при этом свою силу. См.: Horowitz A. Is there a problem in physicalist epiphenomenalism? // *Philosophy and Phenomenological Research* 59: 2 (1999). P. 421-434. Впрочем, все плюсы физикалистского эпифеноменализма обесцениваются несостоятельностью самой теории тождества.

⁴ Нельзя поэтому согласиться с тезисом Д. Льюиса о том, что представление о качественном опыте аналитически влечет за собой представление о его каузальной действительности — см.: Lewis D. An argument for the Identity theory // *Journal of Philosophy* 63: 1 (1966). P. 20.

⁵ См. также: Pauen M., Staudacher A., Walter S. Epiphenomenalism: Dead end or way out? // *Journal of Consciousness Studies* 13: 1-2 (2006). P. 7-19.

⁶ У. Робинсон доказывает, что термин «эпифеноменализм» восходит к критике Гексли У. Джемсу. См.: Robinson W. S. Epiphenomenalism: Entry in the *Stanford Encyclopedia of Philosophy*, 2007.

⁷ Впрочем, некоторые трактуют эпифеноменализм широко — так, что в эту рубрику попадает не только параллелизм, но и окказионализм. См.: Crane T. Mental causation. *TS* 2003. P. 5.

⁸ Н. Кемпбелл доказывает, что Гексли занимал именно такую позицию — см.: Campbell N. What was Huxley's epiphenomenalism? // *Biology and Philosophy* 16 (2001). P. 363-365.

⁹ См., напр.: Stoerig P., Cowey A. Blindsight in man and monkey. *Brain* 120 (1997). P. 535-559; Weiskrantz L. The case of blindsight // *The Blackwell Companion to Consciousness*. P. 175-180.

¹⁰ Хорошо известно о существовании этого феномена — открытого Н. Хамфри, который работал в лаборатории знаменитого Л. Вейскранца (см.: Humphrey N. *Seeing Red: A Study in Consciousness*. Cambridge MA, 2006. P. 43-47), — у обезьян, и многие экспериментальные данные получены именно на основании опытов с ними. Собственно, Хамфри тоже работал с ними.

¹¹ См.: Koch C. *The Quest for Consciousness*. P. 205-229.

¹² По крайней мере, об этом говорит Петра Штёриг, известная именно своими исследованиями в области «слепого зрения» — см.: Blackmore S. *Conversations on consciousness*. P. 216.

¹³ См., его обобщающую работу — Libet B. *Mind Time: The Temporal Factor in Consciousness*. Cambridge MA, 2004. P. 33-89, 123-156.

¹⁴ См., напр.: Herrmann C. S., Pauen M., Min B. K., Busch N.A., Rieger J. W. Eine neue Interpretation von Libets Experimenten aus der Aua-

lyse einer Wahlreaktioiisaufrage // *Bewusstsein. Philosophie, Neurowissenschaften, Ethik*, hrsg. v. C. S. Hermann, M. Pauen, J. W. Rieger und S. Schicktanz. Paderborn, 2005. S. 120-134. Авторы доказывают, что потенциал готовности в мозге, предшествующий некоему действию и осознанию желания совершить его сейчас (т. е. осознанию воления), может быть связан не с подготовкой какого-то конкретного действия, как считает Либет, а с ожиданием действия вообще (хотя ср.: Passingham R. E., Lau H. C. Free choice and the human brain // *Does Consciousness Cause Behavior?* Ed. by S. Pockett, W. P. Banks and S. Gallagher. Cambridge MA, 2006. P. 54) — и это оставляет место для свободы воли. В действительности Либет мог бы и согласиться с этой интерпретацией (ср. след. прим.). но проблема в том, что такая позиция не только не устраняет концептуальные трудности, но и не стыкуется с его более общими результатами, согласно которым сознание запаздывает за осознаваемыми событиями.

¹⁵ Сам Либет, правда, утверждал, что, хотя мы не можем сознательно инициировать наши действия, мы можем накладывать сознательное вето на них — или не накладывать его, давая нечто вроде свободного соизволения на совершение действия. Но очевидно, что это непоследовательная позиция, так как само вето или его отсутствие должны были предрешаться на нейронном уровне — Либет, разумеется, отрицает это (Libet B, *Mind Time*. P. 145-147), но делает это по, так сказать, «идеологическим» соображениям, вступая в конфликт с собственными принципами.

¹⁶ Block N. On a confusion about a function of consciousness // *The Nature of Consciousness*. P. 375-415.

¹⁷ Если в двух словах, то Блок пытается найти эмпирическое подкрепление своим выводам, рассматривая, в частности, эксперименты, показывающие, что наличие той или иной феноменальной данности не всегда означает, что субъект может воспользоваться информацией о ней.

¹⁸ Именно такая схема обосновывается в замечательной книге Д. Вегнера «Иллюзия сознательной воли» — см: Wegner D. M. *The Illusion of Conscious Will*. Cambridge MA, 2003 (2002). Отмечу, правда, что Вегнер допускает смешение осознанной воли, т. е. переживания воления, со свободной волей, или чувством свободы (P. 317-328). Даже если он прав в описании деталей генезиса переживания воления — «Переживание сознательной воли возникает, когда мы заключаем, что наше сознательное намерение было причиной нашего произвольного действия» (P. 62), — это еще не означает, что чувство свободы возникает по тем же причинам. Ср: Wayne T. *Phenomenology and the feeling of doing: Wegner on the conscious will* // *Does Consciousness Cause Behavior?* P. 183.

¹⁹ Именно такой путь — концептуального прояснения этого вопроса — кажется мне наиболее перспективным. Противоположную позицию отстаивает О. Фланаган — см.: Flanagan O. *Conscious inessentialism and the epiphenomenalist suspicion* // *The Nature of Consciousness*. P. 372.

²⁰ См.: Ramachandran V. S. *Mirror neurons and imitation learning as the driving force behind "the great leap forward" in human evolution* // *Edge* (2000).

²¹ Ср.: Armstrong D. M. The Mind-Body Problem. P. 50; Place U. T. Identifying the Mind. P. 104.

²² Ср.: Russell B. Analogy // The Nature of Mind. P. 90.

²³ Hasker W. The Emergent Self. Ithaka, 2001(1999). P. 80.

²⁴ Ср.: Jackson F. Epiphenomenal qualia. P. 46-47; Rieger J. W., Schick Tanz S. «Wenn Du denkst, dass ich denke, dass Du denkst...» // Bewusstsein. Philosophie, Neurowissenschaften, Ethik. S. 41-42.

²⁵ О. Фланеган утверждал, что эволюционный аргумент против эпифеноменализма можно усилить, обратив внимание на то, что сознание присуще существам многих видов. Это означает, что оно было переоткрыто эволюцией много раз, что косвенно свидетельствует о его адаптивной пользе — см.: Flanagan O. Conscious inessentialism and the epiphenomenalist suspicion // The Nature of Consciousness. P. 360. Однако вскоре мы увидим, что этот шаг ничего не дает. Он возможен лишь при допущении наличия сознания у других существ, но само это допущение предполагает отрицание эпифеноменализма (см. далее), что делает дальнейшие аргументы излишними.

²⁶ Ср.: Robinson W. S. Evolution and epiphenomenalism // Journal of Consciousness Studies 14: 11 (2007). P. 27-42.

²⁷ См. также: Staudacher A. Epistemological challenges to qualia-epiphenomenalism // Journal of Consciousness Studies 13: 1-2 (2006). P. 153-175.

²⁸ Попытку опровергнуть Чалмерса предпринял и А. Элитцур. Если Чалмерс прав в том смысле, что его собственные рассуждения о сознании не зависят от существования сознания, т. е. если даже зомби-двойник Чалмерса, Чармлесс, с необходимостью рассуждал бы о сознании, то мы, считает Элитцур, сталкиваемся с противоречием, так как этот зомби не мог бы допускать возможность другого зомби (ведь он с необходимостью рассуждает о существовании сознания), а Чалмерс допускает ее, при том, что поведенчески он должен быть неотличим от Чармлесса (см.: Elitzur A. Consciousness makes a difference: reluctant dualist's confession TS 2005; см. также: Elitzur A. Why quantum mind to begin with: a proof for the incompleteness of the physical account of behavior // Quantum Mind 2007. P. 8-9). Недостаточность этого аргумента Элитцура бросается в глаза: из того, что зомби Чармлесс с необходимостью рассуждает о существовании собственного сознания, не следует, что он с необходимостью должен рассуждать о существовании сознания в других системах, в частности у его физического двойника.

²⁹ Кирк допускает явное противоречие, утверждая, что, с одной стороны, «зомбисты», т. е. те, кто допускает представимость зомби, должны признавать представимость эпифеноменализма, а с другой стороны, доказывая, что, представляя особого зомби (см. основной текст), можно показать непредставимость эпифеноменализма — Kirk R. Zombies mid Consciousness. N.Y., 2005. P. 45-50. Замечу, что подобные нестыковки обесценивают позитивную часть теории Кирка.

³⁰ На русском языке на эту тему см.: Гаспаров И.Г. Представимость зомби и психофизическая проблема // Философия сознания: классика и современность. М., 2007. С. 121-127.

³¹ После того как такие основания будут найдены, мы, как окажется, должны будем говорить, что, хотя зомби логически возможны, мы

не можем верить в их реальное существование. Тезис о том, что зомби могут быть логически невозможны, очевидно ложен, так как в случае его истинности мы не могли бы даже представить, что другие люди лишены квалиа, тогда как, повторю, другие люди непосредственно и отчетливо представляются нами без каких-либо квалиа.

³² Другие вариации подобной критики Чалмерса не выглядят более эффективными — см., напр.: Rivas T., Dongen H. v. Exit epiphenomenalism: the demolition of a refuge // JNL RMI 2:1 (2003).

³³ См., напр.: Moore G. E. *Some Main Problems of Philosophy*. N. Y., 1962 (1953). P 13-41.

³¹ Beattie J. *An Essay on the Nature and Immutability of Truth in Opposition to Sophistry and Scepticism*, 3ed., L., 1772. P. 45.

³⁵ См., напр.: Reid T. *The Works of Thomas Reid*. 7 ed. Edinburgh, 1872. P. 425. При этом интуиция понималась ими, скорее, не в декартовском, а в обыденном смысле.

³⁶ Либо тем, что они отображают единообразный опыт.

³⁷ Образцы такого рода феноменологии встречаются в работах Юма. Но тексты, где она реально практиковалась в последующие времена, надо еще поискать — их довольно мало.

³⁸ Хотя это не означает, что окрестность воспроизводимого события дана так же наглядно, как и само это событие. Скорее, она затенена, и именно поэтому наши яркие воспоминания часто напоминают скорее фотографии, чем видео (факт, осознанию которого я обязан М. Кундере).

³⁹ Иными словами, факты искажения воспоминаний нерелевантны для данного анализа.

⁴⁰ Кстати, формально Юм говорил о «сходстве» прошлого и будущего, но в другом месте я показывал, что *de facto* он также работал с принципом их тождества — см.: Васильев В.В. *История философской психологии. Западная Европа — XVIII век*. Калининград, 2003. С. 204-213.

⁴¹ Под совокупным событием я понимаю примерно то, что Витгенштейн понимал под миром.

⁴² Выделение подобных компонентов является во многом произвольным, и тезис об их наличии еще не означает признание существования каких-либо атомарных событий. Таким образом, последующие рассуждения никак не связаны с принятием определенной онтологии событий.

⁴³ См.: Ramachandran V. S., Hirstein W. Three laws of qualia: What neurology tells us about the biological functions of consciousness, qualia and the self // *Journal of Consciousness Studies* 4: 4-5 (1998). P. 432. Авторы полагают, что так можно было бы ощущать квалиа других существ.

⁴⁴ Но даже этот вариант маловероятен. Он мог бы проходить при условии истинности принципа локальной супервентности, так как в этом случае для воспроизведения приватных состояний другого достаточно было бы воспроизвести сопровождающие их нейронные процессы. Однако в дальнейшем мы увидим, что принцип локальной супервентности, судя по всему, ошибочен.

⁴⁵ Б. Монтеро пробовала найти оптимальную формулу принципа каузальной замкнутости. На ее взгляд, традиционная формулировка «каждый физический феномен имеет физическую причину»

«с одной стороны — слишком слаба, с другой — излишне сильна». Излишне сильна, «так как некоторые физические феномены вообще могут не иметь причин» (в качестве примера она упоминает Большой взрыв), чересчур слаба — потому что «она оставляет открытой возможность того, что для физического действия может быть также необходимой и нефизическая причина» — Montero B. Varieties of causal closure // *Physicalism and Mental Causation*. P. 174. Но выбор формулировки не дело произвола, и мы видели, что анализ наших базовых убеждений приводит именно к традиционной формулировке этого принципа, разве что речь должна идти о «непосредственных причинах», чтобы избежать допущения промежуточных ментальных причин. Ср.: Lowe E.J. Causal closure principles and emergentism // *Philosophy* 75: 4 (2000). P. 571-585.

¹⁶ См., напр.: Chalmers D. J. *The Conscious Mind*. P. XIII.

⁴⁷ См.: Papineau D. *Thinking about Consciousness*. P. 232-256. Ср.: Montero B. Varieties of causal closure // *Physicalism and Mental Causation*. P. 177-180.

⁴⁸ Как, к примеру, считает У. Майкснер — см.: Meixner V. New perspectives for a dualistic conception of mental causation // *Journal of Consciousness Studies* 15: 1 (2008). P. 17.

⁴⁹ Сходные — но лишь на первый взгляд — тезисы высказывает Вим де Муижнк в монографии «Зависимости, связи и другие отношения: теория ментальной каузальности» (2003). В частности, пытаясь преодолеть эпифеноменализм, он утверждает, что ментальные состояния относятся к числу так называемых «реляционных свойств», формирующихся в том числе на основе прошлого опыта. Таким образом, наличие ментальных состояний позволяет организму учитывать свою историю, и это делает их каузально действенными. Между тем — и это, на мой взгляд, является его главным просчетом — Муижнк отделяет вопрос о каузальной действенности ментальных состояний от вопроса о каузальной действенности квалиа, подчеркивая, что не в состоянии ответить на второй из этих вопросов — см.: Muljnck W. *Dependencies, Connections, and Other Relations: A Theory of Mental Causation*. Dordrecht, 2003. P. 256-257. Но тогда он должен был бы понимать ментальные состояния в механистическо-функционалистском смысле. Система, реализующая такие состояния, «учитывала» бы свою историю по большому счету так же, как «учитывает» ее любой чисто физический объект, т. е. никак в сравнении с человеком. В действительности Муижнк отвергает и традиционную функционалистскую трактовку ментальных состояний, выдвигая «альтернативный взгляд», согласно которому эти состояния существуют на стыке организма и его окружения и в таком подвешенном модусе служат «ограничителями» его каузальных цепей — теория с марксистским оттенком, не допускающая ясного толкования. См.: Muijnck W. Two types of mental causation // *Philosophical Explorations*, 7:1 (2004). P. 21-35.

⁵⁰ К примеру, бессмысленно пытаться критиковать ее ссылкой на невозможность точного повторения событий. Ведь понятие события относительно, а какие-то аспекты происходящего совершенно точно повторяются с другими причинами. Их и можно брать за осно-

ву. Конечно, повторение происходит на фоне различий, но это никак не противоречит тому, что я утверждаю.

⁵¹ О значении модели Конвея — см.: Dennett D. C. *Freedom Evolves*. L., 2004 (2003). P. 36-47.

⁵² «Я считаю, что если разные истории приводят к двум идентичным мозгам, то у этих мозгов будет опыт одного и того же рода -различные истории окажутся нерелевантными для характера их опыта» — письмо от 25 октября 2004 г. Из другого письма (14 ноября 2004 г.): «Конечно, я считаю естественно возможной (*naturally possible*) такую ситуацию, когда различные истории приводят к одному и тому же состоянию мозга. Но если это происходит, то, на мой взгляд, у этих мозгов будут одинаковые воспоминания — различия в их прошлом не будут отражаться в различных воспоминаниях (если бы они отражались, то мы имели бы разные состояния мозга)».

⁵³ Ср.: Васильев В.В. Мозг и сознание: выходы из лабиринта // Вопросы философии. № 1, 2006. С. 67-79. К сходным выводам, но на основе других аргументов и мысленных экспериментов пришел Ю. Фишер. См.: Fisher J, *Why nothing mental is just in the head* // *Nous* 41 (2007). P. 318-334.

⁵⁴ Именно поэтому нельзя использовать данный тезис как довод против этого аргумента.

⁵⁵ Ср.: Kim J. *Philosophy of Mind*. Cambridge MA, 2006. P. 8-11. Ким предлагает несколько формулировок принципа супервентности: (1) «Ментальное супервентно на физическом в том смысле, что вещи (объекты, события, организмы, личности и т. п.) совершенно одинаковые в своих физических свойствах не могут отличаться относительно ментальных свойств. Иными словами, физическая неотличимость влечет психологическую неотличимость» (локальная супервентность), (2) «Ментальное супервентно на физическом лишь в том случае, когда, если нечто X имеет ментальное свойство M, то имеется физическое свойство P, такое, что X имеет P, и необходимо, что всякий объект, обладающий P, наделен M», (3) «Ментальное супервентно на физическом в том смысле, что миры, одинаковые во всех физических отношениях, одинаковы и во всех ментальных отношениях; т. е. миры, одинаковые физически, совершенно одинаковы и в целом» — *idid*. P. 9-10 Последняя формулировка выражает так называемую «глобальную супервентность», и мы увидим, что предложенные мною аргументы не отрицают ее. Вторая формулировка тоже не опровергается ими, если она допускает глобальную интерпретацию. Ошибочно лишь первое понимание. Отмечу в этой связи, что предложенная трактовка не означает поддержку «аномального монизма» Д. Дэвидсона. Ведь, помимо того, что он рассматривал его как вариацию теории тождества, он но отрицал принципа супервентности во всех его формах — см.: Davidson D. *Mental events* // *The Nature of Mind*. P. 250, О соотношении различных формулировок принципа супервентности — см., напр.: Baker L. R. *Metaphysics and mental causation* // *Mental Causation*, ed. by J. Heiland A. Mele. N. Y., 1993. P.79-80.

⁵⁶ Неверно было бы оспаривать корректность этого сценария на основании того, что существование абсолютно одинаковых индивидов является крайне маловероятным. Ведь для действительности пред-

ложенного эксперимента достаточно лишь, чтобы все это было реально возможным. В этом плане данный мысленный эксперимент не отличается от множества аналогов.

⁵⁷ Письмо от 18 ноября 2004 г.

⁵⁸ Hofstadter D., Dennett D. C. *The Mind's I: Fantasies and Reflections on Self and Soul*. N.Y., 2000 (1981). P. 3-4, 7 («Глаз разума» — авторизованное (в этом нет сомнений) заглавие русского перевода этой книги). См. также: Parfit D. *Reasons and Persons*. N. Y., 1986 (1984). P. 199-201.

⁵⁹ Ср.: Clarke C. J. S. *The nonlocality of mind // Explaining Consciousness*. P. 165-175. Автор совершенно другим путем приходит к выводу о связи физической и ментальной нелокальности.

⁶⁰ Тонкость в том, что контрфактуальная теория причинности, которую я поддерживаю — хотя я и подхожу к ней несколько иначе, чем другие ее сторонники, прежде всего Д. Льюис (см., например, его статью «Причинность» (1973) — Lewis D. *Philosophical Papers*. V. II. N. Y. 1986. P. 159-172; в частности, я не считаю перспективным работать с обыденным понятием причинности — акцентирование которого приводит, помимо прочего, к парадоксам сверхдетерминации — и, кроме того, не принимаю без уточнений то близкое Льюису контрфактуальное определение причинности, которое восходит к Юму и так связывает два «объекта», причину и действие, что «если бы первого не было, второй никогда бы не существовал»: Юм был уверен, что у события может быть только одна причина, но никак не обосновывал это), — запрещает мыслить такое уничтожение. В самом деле, отношение физической причинности можно задать следующим контрфактуальным определением: причина некоего физического события В — это такое физическое событие А, которое является элементом класса событий 2, такого, что каждый из его элементов обладает тем свойством, что его отсутствие в мире близком актуальному миру настолько, что можно говорить, что в нем действуют те же физические законы (а его состояния можно рассматривать в качестве будущих состояний актуального мира), при отсутствии на его месте других элементов этого класса — привело бы к тому, что В не произошло. В ином случае В происходит. Иначе говоря, причина события В — это такое событие А, что отсутствие события В в мире с теми же физическими законами означало бы отсутствие события А. Посмотрим, применимо ли это к нашему случаю. Я верю, что мое поведение скоррелировано с каким-то нелокальным физическим событием или совокупностью событий А. Т. е. я убежден, что если бы я — будучи таким же физически — вел себя иначе, то моему измененному поведению сопутствовал бы другой нелокальный коррелят. Значит, это А может считаться причиной моего поведения.

⁶¹ Эта конструкция может служить иллюстрацией того, что Дж. Лау (рассуждая, впрочем, лишь в проблематическом ключе) назвал «причинением ментальным событием физического каузального факта». См.: Lowe E. J. *Causal Closure Principles and Emergentism*. P.571-585.

⁶² Этим, стало быть, завершается феноменологическая дедукция понятия ментальной каузальности.

⁶³ Обзор интерпретаций квантовой механики и их отношения к каузальной роли сознания — см.: Rosenblum B., Kuttner F. *Quantum Enigma: Physics Encounters Consciousness*. N. Y., 2006.

⁶¹ См., напр.: Tegmark M. The importance of quantum decoherence in brain processes // *Physical Review E* 61 (2000). P. 4194-4206; Hagan B., Hameroff S., and Tuszynski J. Quantum computation in brain microtubules: decoherence and biological feasibility // *Physical Review E* 65 061901 (2002).

⁶⁵ Можно было бы порассуждать о том, что предложенная концепция сочетается с тезисом об эквивалентности Копенгагенской интерпретации и интерпретации скрытых параметров Д. Бома.

⁶⁶ Ср.: Hameroff S. Quantum computation in brain microtubules? The Penrose-Hameroff 'Orch OR' model of consciousness // *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* 356: 1743 (1998). P. 1869-1896.

⁶⁷ Ср.: Zizek S. *The Parallax View*. Cambridge MA, 2006. P. 240.

⁶⁸ См., напр.: Wolff Chr. *Vernunfftige Gedanken von Gott, der Welt und der Seele des Menschen, auch alien Dingen uberhaupt*. 3 Aufl. Halle, 1725. S. 454. Наиболее проработанную трактовку сознания как способности различения предложил в XVIII веке И. Н. Тетенс в трактате «Философские опыты о человеческой природе и ее развитии» — см.: Tetens J. N. *Philosophische Versuche uber die menschliche Natur und ihre Entwicklung*. Bd. 1 — 2 Lpz., 1777. Bd. 1. S. 262 ff. В современной отечественной литературе тему сознания как различения развивает феноменолог В. И. Молчанов. См.: Молчанов В.И. *Различение и опыт: феноменология неагрессивного сознания*. М., 2004.

⁶⁹ Это допущение позволяет отрицать возможность раздвоения субъекта — к примеру, при мысленном или реальном разделении полушарий мозга (в противном случае мы можем прийти к парадоксам — ср.: Parfit D. *Personal identity* // *Personal Identity*, ed. by J. Perry. Berkeley, 1975. P. 199-203). Но подобное допущение не отменяет возможности раздвоения ментальной жизни. Ср.: Nagel T. *Brain bisection and the unity of consciousness* // *Synthese* 22 (1971). P. 396-413.

⁷⁰ Суть этой теории в одной фразе: «Ментальное состояние сознательно, если оно сопровождается соответствующей мыслью более высокого порядка» — Rosenthal D. *Two concepts of consciousness* // *The Nature of Mind*. P. 465. Ментальное состояние более высокого порядка может быть и бессознательным (если оно все же осознается, то у нас имеется мысль о мысли о ментальном состоянии, т. е. это случай самосознания или интроспекции). Поэтому Розенталь противопоставляет свое понимание картезианской концепции, подразумевающей отождествление ментальных и сознательных состояний. См. также: Rosenthal D. *A Theory of consciousness* // *The Nature of Consciousness*. P. 729-753; Carruthers P. *Higher-order theories of consciousness* // *The Blackwell Companion to Consciousness*. P. 277-286. Альтернативные подходы — см.: *Self-Representational Approaches to Consciousness*, ed. by U. Kriegel and K. Williford. Cambridge MA, 2006.

⁷¹ Разумеется, иллюзорность чувства свободы не означает иллюзорность морали и т. и. Это разные вещи, и если говорить о моральных принципах и их истоках, то в этом вопросе за последние десятилетия

было сделано так много, что его можно, на мой взгляд, считать почти решенным. После работ Р. Доукинса и ряда других авторов стало ясно, что альтруистические чувства, существующие у людей, прекрасно объясняются с точки зрения теории эволюции, так как они способствуют реплицированию генетического материала. Правда, уже давно было показано, что моральное чувство не всегда может оценивать как моральное то, что продиктовано альтруистическими побуждениями. Истина, возможно, в том, чтобы рассматривать его как интегральное чувство, позволяющее установить оптимальный баланс между альтруизмом и эгоизмом. Особенно симпатична мне та версия этой теории, которая была выдвинута еще в 1790 г. автором немецкой книги «Наблюдения о человеческом духе и его отношении к миру», скрывавшимся под псевдонимом Андрей П. Колыванов (настоящее имя автора, возможно, А. И. Вяземский). Из современных теоретиков, на мой взгляд, лучше всего об этом пишет Пинкер. Напротив, П. Сингер с его громкими призывами разрушить старые нравственные скрижали и внедрить новую гипертруистическую этику, даже если ради этого придется поработать с нашими хромосомами, кажется мне столь же далеким от истины, как и Ницше — его антипод столетней выдержки.

⁷² Разумеется, речь идет не о выживании индивида, а о репликации ДНК.

⁷³ Хочу оговориться, что этот тезис не имеет отношения к традиционной экстерналистской позиции (классической иллюстрацией которой является мысленный эксперимент «двойник Земли» Х. Патнэма). Экстерналисты утверждают, что одни и те же супервентные внутренние состояния, т. е. состояния, в полной мере детерминированные физическим состоянием мозга, могут отсылать к разным референтам, и эти различия зависят от разного рода внешних факторов. Я же утверждаю нечто иное, а именно что сам внутренний контент ментальных состояний не является производным от состояния мозга, так как зависит еще и от окружения и истории организма.

⁷⁴ См., напр.: Tuszynski J. (ed.) *The Emerging Physics of Consciousness*. В., 2006.

⁷⁵ Одновременно с «трудной» решается и так называемая «более трудная проблема сознания» (термин Н. Блока): «Почему физически различные существа должны иметь нечто общее в феноменальном плане именно так, а не иначе или вообще никак?» (Block N. *The harder problem of consciousness*. P. 423), Речь у Блока идет о том, на каком основании мы можем признавать или отрицать сознание у существ, демонстрирующих сходное с нами поведение, но физически совершенно непохожих на нас. Ответ таков: при отсутствии сознания другие существа не могли бы демонстрировать сходное с нами поведение, а именно проходить тотальный тест Тьюринга.

⁷⁶ Этот термин использовался и раньше — в совершенно другом смысле. В частности, его применял Р. Сперри, который понимал эмерджентные свойства, благодаря которым осуществляется «нисходящая» причинность, по сути, как высокоуровневые описания состояний и процессов, существующих на базовом физическом уровне, так как подчеркивал, что они «не нарушают и не вмешиваются в каузальные отношения», имеющиеся на этом базовом уровне, — Sperry R. *In defense of mentalism and emergent interaction* // *The*

Journal of Mind and Behavior 12: 2(1991). P. 230. Еще более далеко от представленной в данной книге концепции толкование «эмерджентного интеракционизма» Ч. Тартом. Однако, поскольку этот термин не получил широкого распространения, я счел возможным оставить его в качестве альтернативного именованья.

⁷⁷ См.: Sperry R. W. Hemisphere decoimection and unity in conscious awareness // American Psychologist 23 (1968). P. 723-733; Col-oin M., Gazzaniga M. Split-brain cases // The Blackwell Companion to Consciousness. P. 181-193.

⁷⁸ Popper K., Eccles J. The Self and Its Brain: An Agrument for In-teractionism. B., 1977. P. 565-566.

⁷⁹ См.: Kim J. Physicalism, or Something Near Enough. Princeton, 2005. P. 70-85.

⁸⁰ Впрочем, Дж. Фостер, который одним из первых заговорил о проблеме парности (см. также: Sosa E. Mind-body interaction and supervenient causation // Midwest Studies in Philosophy 9 (1984). P. 271-281), показывает, что картезианский дуализм мог бы справляться с ней. См.: Foster J. The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind. L., 1996 (1991). P. 162-172.

⁸¹ Подобным образом можно объяснить и само порождение этих (а не каких-либо иных) состояний, если задуматься о решении проблемы установления пар от физического к ментальному.

⁸² Б частности, Хескер показывает, что используемая Сперри иллюстрация эмерджентной нисходящей каузальности - колесо, движение которого обусловлено его макроскопическими свойствами, — не решает проблему интеракционизма. См.: Hasker W. The Emergent Self. P. 178-188.

⁸³ Много конкретики можно найти в книге Д. Хадгсона «Сознание значимо» (1991). Как и Хескер, Хадгсон отстаивает интеракционизм, но проблема его книги в том, что автор считает вопрос о ментальной каузальности и другие смежные вопросы эмпирическими, а не концептуальными — см.: Hodgson D. The Mind Matters: Consciousness and Choice in a Quantum World. N. Y., 1991. P. 3. Такой подход крайне осложняет задачу их решения и создает впечатление незавершенности многих аргументов, а также размытости тезисов, представленных в этой работе.

⁸⁴ См.: Levins J, Purple Haze: The Puzzle of Consciousness. N. Y., 2004(2001). P. 76-78.

⁸⁵ Нет необходимости также утверждать, как это делает, к примеру, У. Робинсон (см.: Robinson W. S. The hardness of the hard problem // Explaining Consciousness P. 160; сходные идеи высказывал и Т. Нагель), что «трудная проблема» могла бы быть решена лишь при кардинальном изменении наших концептуальных схем.

⁸⁶ Ср.: Hodgson D. The easy problems ain't so easy // Explaining Consciousness. P. 125-131; Lowe E. J. There are no easy problems of consciousness // Explaining Consciousness. P. 117-123,

⁸⁷ Hawkins J., Blakeslee S. On Intelligence. N. Y., 2004.

⁸⁸ Bennett M. et al. Neuroscience and Philosophy: Brain, Mind, and Language. N.Y., 2007.

ПОСЛЕСЛОВИЕ

Мне остается сказать лишь несколько слов о характере проведенных разысканий. Не буду много рассуждать о первых главах, в которых был представлен обзор современных теорий сознания, опирающийся на традиционные историко-философские методы — с тем лишь отличием, что рассматриваемые системы еще не вполне стали историей и ушли в прошлое: это, так сказать, история философии в реальном времени — увлекательная вещь, мечта всех историков мысли, когда можно прямо поинтересоваться у философа, которым занимаешься, хотел ли он сказать то, что, как мы полагаем, он сказал. И конечно, весьма интригующей для меня была работа с рукописями Чалмерса, философа, которому только-только исполнилось 40 лет.

Хотя Чалмерс и другие наши главные герои не кажутся мне оборванными на полуслове. Их идеи уже выстроились в цельные картины, что и позволяет говорить о них в историческом ключе. Картины эти, на мой взгляд, интересны сами по себе — в конце концов, речь идет о ведущих философах начала XXI столетия, — и поэтому я пытался рассказать о них с достаточной детализацией. Но они интересны и как выражения взаимодополняющих подходов к решению проблемы сознания. Поэтому, хотя нельзя исключать, что в позициях Сёрла, Деннета и Чалмерса, которым были посвящены «исторические» главы, произойдут какие-то подвижки¹, но если анализ был верен, то они могут свестись либо к тому, что они просто обменяются взглядами — и тогда мы не увидим прогресса в решении проблемы сознания, — либо они могут стать движением в направлении той концепции, которая была развернута в последней главе.

Представляется, что аргументы, изложенные в этой последней главе и подводющие к нарисованной там картине соотношения мозга и сознания, можно классифицировать в качестве вариаций на тему аналитической философии. Правда, это феноменологизированный вариант аналитической философии, и необходимость такой трансформации² я видел в том, что иначе анализ мог бы остаться на лингвистическом уровне (аналитическая философия исторически была тесно связана с разбором лингвистических конструкций³), а такой уровень, как показывает опыт, недостаточен для решения «трудной проблемы». Отмечу также, что высказанные там тезисы неверно трактовать как гипотезы, надстроенные над эмпирическим материалом.

Я подчеркиваю это потому, что многие аналитические философы в наши дни указывают на гипотетичность собственных построений. Причем иногда они предлагают понимать свои выводы в таком смысле, что по сути те оказываются мало в чем отличными от стандартных экспериментальных гипотез, характеризуясь разве что гораздо меньшей конкретностью. В этом можно усмотреть влияние Куайна, отступившего в свое время от стандартов классической аналитической философии и взявшего курс на размывание границ между философией и эмпирической наукой. Но я по-прежнему убежден в перспективности концептуального анализа⁴. И в целом я рассматриваю свои выводы как результаты строгих дедукций, хотя в ряде пограничных областей эти выводы все же порой приобретали конъектурный характер. Но природа этих дедукций, а также принципов, на которых они базировались, нуждается в дополнительном уточнении. Истоком многих заключений было основоположение о соответствии прошлого и будущего, которое является фундаментом нашего эмпирического познания. Поскольку это основоположение не необходимая истина, то все тезисы, рассматривавшиеся в последней главе, специфичны: они говорят о том, как мы должны представлять отношение сознания и мозга при наличном устройстве наших когнитивных способностей, в какое положение вещей мы должны верить. Ничего большего там нет и быть в принципе не могло.

И здесь уместно еще раз вспомнить об одной примечательной философской традиции, заложенной в Новое время Юмом и позже продолженной Авенариусом и Гуссерлем. Некоторых аналитических философов, в частности Мура и Стросона, тоже можно отнести к ней. Я имею в виду традицию понимания философии как предприятия, главной задачей которого является прояснение обыденных интуиций человеческого сознания. Каждый из нас является носителем некоего стихийного мировоззрения. В нем присутствуют самые разные компоненты. Некоторые из них случайны, другие повторяются у всех людей, кем бы они ни были. Эти универсальные моменты человеческого мировоззрения, вырастающие из базовых онтологических установок — или естественных установок, как сказал бы Гуссерль, — не только образуют основу здравого смысла в его когнитивном аспекте, но и являются предметом особого интереса для философов, которые могут попытаться отрефлексировать эти исходные убеждения.

Но такая рефлексия — очень непростое дело. Уже Юм заметил, что при ее небрежном проведении мы быстро приходим к парадоксам. И тогда возникает странное чувство опасности философии. В самом деле, пока мы не философствуем,

а просто отдаемся во власть естественных установок, мир кажется понятным и логичным. Но стоит вдуматься в них, как перед нами вырастает угроза тотального скептицизма, из которого только один путь — в никуда.

Так вот, искусство философа, по-видимому, состоит в том, чтобы так прояснить естественные установки, чтобы не войти в противоречие с их изначальными предписаниями, чтобы итоговый концептуальный образ мира не расходился с его наивным изображением.

Иначе говоря, философ, как мне кажется, должен ставить перед собой цель защитить наивные установки нашего сознания, подтвердить непротиворечивость предписаний здравого смысла. Внимательные читатели, наверное, заметили, что именно этой цели я пытался достичь на протяжении всей работы. И теперь настало время, когда можно задуматься над тем, оказался ли этот план реализован.

Насколько концепция «локального интеракционизма», развернутая в пятой главе книги, соответствует установкам здравого смысла — вот вопрос, который мы сейчас кратко обсудим. Кратко, потому что многие акценты уже были проставлены в основном тексте. В частности, мы видели, что указанная концепция согласуется с такими положениями здравого смысла, как убеждение в том, что наши желания и другие внутренние состояния влияют на поведение, что другие люди обладают сознаниями и что все события, происходящие в физическом мире, имеют физические же причины,

Но здравый смысл предписывает нам и кое-что еще. Сознывая мир в опыте, мы верим, что непосредственно контактируем с вещами, мы думаем, что видим сам мир, а не его ментальный отпечаток, не его отражение в нашем мозге. Между тем многие философы, рассуждавшие о специфике человеческого восприятия, приходили к выводу, что в восприятии нам даны не сами вещи, а лишь их образы.

Может ли предложенная концепция спасти убежденность здравого смысла в непосредственном контакте нашего Я с вещами? Похоже, что да, по крайней мере потенциал для этого у нее есть. Вспомним, что, согласно полученным ранее выводам, зная наличное физическое состояние мозга, мы не можем с полной уверенностью сказать, что с ним связаны именно такие, а не иные квалиа, приватные ментальные состояния, даже если этот мозг является точной копией моего — а я знаю, какими квалиа я в данный момент наделен.

Какое отношение этот вывод имеет к теории восприятия? Самое прямое. Ведь если мой мозг, взятый в своем физическом аспекте, может быть носителем разных ментальных контентов, но при этом фактически сопряжен со вполне опреде-

ленным их набором — в который входят и непосредственные данные чувств, — то возникает вопрос, почему он обладает именно таким набором? Мы уже знаем, где надо искать ответ на него: в истории его взаимодействия со своим окружением. Но как надо мыслить отношение подобных ментальных контентов и самого мозга? Хотя работающий мозг создает условия для их появления, это отношение нельзя мыслить по модели каузального порождения их мозгом: если бы это было так, то одинаковый мозг всегда порождал бы одинаковые ментальные контенты. Значит, надо перевернуть каузальную нить и рассматривать контенты, среди которых, напомним, есть и чувственные данные, как нечто причиняющее определенные состояния мозга. Иными словами, в соответствии с предложенной концепцией, данные чувств, непосредственно осознаваемые мной, должны рассматриваться как причины, вызывающие определенные изменения моих органов чувств и мозга.

Но ведь это полностью гармонирует с установками здравого смысла, предписывающего нам смотреть на данные чувств как на предметы, т. е. как на сущности, воздействующие на нас, и поэтому как на то, что находится вне нашего организма, а не как на то, что существует в мозге вследствие активации таких-то групп нейронов. Так что тезисы, выдвинутые в данной книге, позволяют защитить концепцию непосредственного восприятия, диктуемую здравым смыслом. Здесь, конечно, остается немало проблем, и я не претендую на изложение цельной и развернутой теории: для меня важно было всего лишь подчеркнуть, что согласие предложенной трактовки отношения сознания и мозга со здравым смыслом простирается за пределы тех границ, которые были очерчены в основной части.

Защита здравого смысла как одна из главных задач философии — это, как представляется, довольно возвышенная цель. Следование здравому смыслу отличает нормальных людей, а мне всегда казалось, что философы — это самые нормальные люди на земле. И нормальные люди живут в реальном, а не в выдуманном ими мире.

Расставание с вымыслами — путь к нормальной жизни. Но этот путь нельзя пройти без сражений с метафизикой, если понимать под ней концептуальную экстравагантность и увлеченность тем, что находится за пределами повседневного. Долгая история философии оставила нам самые разные толкования ее предмета и ее конечного назначения. В XX веке многие считали, к примеру, что философы должны стремиться к демонтажу и элиминации метафизики (это, конечно же, кантианский мотив). В каком-то смысле я говорю о том же.

Истинная философия помогает распрощаться с фантастическими конструкциями, выдаваемыми за реальность, и обратить нас к самой реальности — со всеми ее красотами и несообразностями.

Освобождая фантазию от конструирования реальности, философия создает предпосылки для ее продвижения в деле продуцирования образов и идей, к которому та изначально предуготована. Иначе говоря, она открывает дорогу для подлинной фантазии, помогающей философам — не меньше, чем писателям, — создавать свои тексты. Мне трудно сказать, насколько это справедливо для данной книги. Некоторое время назад ценимая мною начальница одного из наших крупнейших философских издательств поинтересовалась у знающего меня человека, зачем я пишу эту работу, ведь никаких явных целей у меня нет, докторская диссертация защищена и т. п. Этот примечательный вопрос заставил меня всерьез задуматься. И я не нашел лучшего ответа, как — просто для развлечения. Написание этой книги было радостью для меня, и даже жаль, что она закончена.

Примечания

¹ С другой стороны, нельзя исключать, что «подвижки» могут произойти и в позиции, к примеру, Аристотеля, если вдруг отыщутся папирусы с его давно утраченными диалогами. Разумеется, вероятность такой находки ничтожно мала. Но, вообще говоря, переосмысление идей философов прошлого в связи с находкой или изданием их неизвестных или неосвоенных текстов — рутинная вещь, и обнародование старых текстов, по сути, мало отличается от публикации новых.

² Из классиков аналитической философии ближе всех к такому подходу был, пожалуй, П. Стросон. Но из классиков аналитической философии именно он менее всего повлиял на меня.

³ Вместе с тем нет оснований утверждать, как это делал М. Даммит, что философия является «аналитической» тогда и только тогда, когда она занимается анализом языка — см.: Dummett M. *Origins of Analytical Philosophy*. P. 4. Это совершенно произвольное ограничение.

⁴ Ср.: Jackson F. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. N. Y., 1998.

Литература

- Алексеев А. Л. Философский текст: идеи, аргументация, образы. М., 2006.
- Армстронг Д.М. Льюис и теория тождества // Историко-философский альманах: Вып. 2. М., 2007. С. 227-231.
- Блинов А. Л. и др. Аналитическая философия. М., 2006.
- Блохина Н.А. Теория тождества в системе современных представлений о сознании // Философия сознания: история и современность. М., 2003. С. 181-188.
- Вальверде К. Философская антропология. М., 2000.
- Васильев В.В. История философской психологии. Западная Европа — XVIII век. Калининград, 2003.
- Васильев В.В. Мозг и сознание: выходы из лабиринта // Вопросы философии. 2006. № 1. С. 67-79.
- Васильев В.В. Кока-кола и секрет Китайской комнаты // Философия сознания: классика и современность. М., 2007. С. 86-94.
- Величковский Б.М. Когнитивная наука. Т. 1-2. М., 2006.
- Веретенников А.Л. Функционализм и философия тождества // Философия сознания: классика и современность. М., 2007. С. 95-104.
- Вильянуэва Э. Что такое психологические свойства? Метафизика психологии. М., 2006.
- Витгенштейн Л. Философские работы. Ч. 1. М., 1994.
- Волков Д.Б. «Сильная версия искусственного интеллекта» // Философия сознания: классика и современность. М., 2007. С. 105-113.
- Гарнцев МЛ. Проблема самосознания в западноевропейской философии (от Аристотеля до Декарта). М., 1987.
- Гарнцева Н.М. Сознание как комплекс мемов // Философия сознания: классика и современность. М., 2007. С. 114-120.
- Гаспаров И.Г. Представимость зомби и психофизическая проблема // Философия сознания: классика и современность. М., 2007. С. 121-127.
- Гаспарян Д.Э. О физическом и логическом типах отношений между сознанием и телом // Философия сознания: классика и современность. М., 2007. С. 128-133.
- Грязное А.Ф. Аналитическая философия. М., 2006.
- Гуссерль Э. Идеи к чистой феноменологии и феноменологической философии. Кн. 1. М., 1999.
- Декарт Р. Сочинения в 2 т. Т. 2. М., 1994.
- Деннет Д. Виды психики: на пути к пониманию сознания. М., 2004.
- Доброхотов А.Л. Категория бытия в классической западноевропейской философии. М., 1986.
- Доброхотов А.Л. Избранное. М., 2008.
- Дубровский Д.И. Психические явления и мозг. Философский анализ проблемы в связи с некоторыми актуальными задачами нейрофизиологии, психологии и кибернетики. М., 1971.

- Дубровский Д.И. Сознание, мозг, искусственный интеллект. М., 2007.
- Дэвидсон Д. Истина и интерпретация. М., 2003.
- Жучков В.А. Кант и проблема сознания // *Философия сознания: история и современность*. М., 2003, С. 52-74.
- Золкин А.Л. Гипотеза «языка мышления» // *Философия сознания: классика и современность*. М., 2007. С. 134-139.
- Зотов А.Ф. Логико-семантические предпосылки исследовательской программы искусственного интеллекта // *Философия сознания: классика и современность*. М., 2007. С. 314-323.
- Иваницкий А.М. Мозговая основа субъективных переживаний: гипотеза информационного синтеза // *Журнал высшей нервной деятельности им. И.П. Павлова*. 1996. № 46:2. С. 241-252.
- Иванов А.В., Миронов В.В. Университетские лекции по метафизике. М., 2004.
- Иванов Д.В. Дуализм и qualia // *Вестник Московского университета. Сер. Философия*. 2006. № 2. С. 3-21.
- Иванов Д.В. Проблема сознания и философия Витгенштейна // *Вестник Московского университета. Сер. Философия*. 2007. №2. С. 6-19.
- Косилова Е.В. Изучение сознания в психологии // *Философия сознания: классика и современность*. М., 2007. С. 324-330.
- Крафт В. Венский кружок. М., 2003.
- Кротов АЛ. Философия Мальбранша. М., 2003.
- Кузнецов В.Н. Проблема сознания в философии Жюльена-Офре де Ламетри и «философия сознания» Джона Сёрла // *Философия сознания: история и современность*. М., 2003. С. 28-35.
- Кузнецов В.Н. Европейская философия XVIII века. М., 2006.
- Кузнецов В.Н. Проблема «сознание—тело» в философии Нового времени // *Философия сознания: классика и современность*. М., 2007. С.18-24.
- Лекторский В.А. Возможна ли интеграция естественных наук и наук о человеке? // *Вопросы философии*. 2004. № 3. С. 44-49.
- Майоров Г.Г. Философия как искание Абсолюта: Опыты теоретические и исторические. М., 2004.
- Макеева Л.Б. Философия Х. Патнэма. М., 1996.
- Мареева Е.В. Дж. Сёрл: старое и новое в понятии сознания // *Философия сознания: история и современность*. М., 2003. С. 216-231.
- Мартынов К.К. Интенциональность как единое проблемное поле аналитической философии сознания и феноменологии. Диссертация на соискание ученой степени кандидата философских наук. М., 2006.
- Молчанов В.И. Различение и опыт: феноменология неагрессивного сознания. М., 2004.
- Мотрошилова Н.В. «Идеи I» Эдмунда Гуссерля как введение в феноменологию. М., 2003.
- Мур Д.Э. Доказательство внешнего мира // *Аналитическая философия: избранные тексты*. М., 1993. С. 66-84.
- Мур Д.Э. Защита здравого смысла//*Аналитическая философия: становление и развитие (антология)*. М., 1998. С. 130-154.

- Остин Дж. Избранное. М., 1999.
- Остин Дж. Три способа пролить чернила: философские работы. СПб., 2006.
- Палмер Дж., Палмер Л. Эволюционная психология. М., 2003.
- Пассмор Дж. Сто лет философии. М., 1998.
- Пассмор Дж. Современные философы. М., 2002.
- Патнэм Х. Философия сознания. М., 1998.
- Патнэм Х. «Гипотеза врожденности» и объяснительные модели в лингвистике // Философия языка. М., 2004. С. 178-190.
- Патнэм Х. Химера когнитивной науки // Философия и будущее цивилизации: Тезисы докладов и выступлений IV Российского философского конгресса (Москва, 24-28 мая 2005 г.). М., 2005. Т. 1. С. 276.
- Пенроуз Р. Новый ум короля: О компьютерах, мышлении и законах физики. М., 2003.
- Пенроуз Р. Тени разума. Ч. 1-2. М., 2003, 2005.
- Петрунин Ю.Ю. Искусственный интеллект: история, методология, философия. М., 2002.
- Пинкер С. Язык как инстинкт. М., 2004.
- Поппер К. Критерий эмпирического характера теоретических систем // Журнал «Erkenntnis» («Познание»). Избранное. М., 2006. С. 335-337.
- Прист С. Теории сознания. М., 2000.
- Райл Г. Понятие сознания. М., 1999.
- Рамачандран В.С. Рождение разума. Загадки нашего сознания. М., 2006.
- Рассел Б. Человеческое познание, его сфера и границы. Киев, 1997.
- Рассел Б. Избранные труды. Новосибирск, 2007.
- Рорти Р. Философия и зеркало природы. Новосибирск, 1997.
- Рорти Р. Бум в исследованиях сознания в свете элиминативизма // Историко-философский альманах: Вып. 2. М., 2007. С. 242-244.
- Руднев В.П. (ред.) Людвиг Витгенштейн: человек и мыслитель. М., 1993.
- Сакс О. Человек, который принял свою жену за шляпу. М., 2005.
- Сёрл Дж. Открывая сознание заново. М., 2002.
- Сёрл Дж. (ред.) Философия языка. М., 2004.
- Сёрл Дж. Рациональность в действии. М., 2004.
- Соколов В.В. Историческое введение в философию. М., 2005.
- Сокулер З.А. Как избавиться от власти картезианской парадигмы? // Философия сознания: классика и современность. М., 2007. С. 173-185.
- Стивенсон Л. Десять теорий о природе человека. М., 2004. С. 187-203.
- Стрельцова Г.Я. Сознание как проблема и тайна (сознание как «жертва» и «жертвы сознания») // Философия сознания: история и современность. М., 2003. С. 321-331.
- Трушкина Н.Ю. Отношение теоретических концепций и компьютерных моделей в исследованиях искусственного интеллекта.

Диссертация на соискание ученой степени кандидата философских наук. М., 2008.

Хомский Н. Картезианская лингвистика. М., 2005.

Хомский. Н. О природе и языке. М., 2005.

Хофштадтер Д. Гёдель, Эшер, Бах: Эта бесконечная гирлянда. Самара, 2001.

Хофштадтер Д., Деннетт Д. (ред.) Глаз разума. Самара, 2003.

Черняк А.З. К вопросу о каузальной эффективности ментальных свойств // Философия сознания: классика и современность. М., 2007. С. 105-113.

Шохин В.К. Психология, метафизика и немного компаративистики (о книге В. В. Васильева «История философской психологии. Западная Европа — XVIII век») // Вопросы философии. 2005. № 2. С. 158-165.

Шохин В.К. Философия ценностей и ранняя аксиологическая мысль. М., 2006.

Эдмондс Д., Айдиноу Дж. Кочерга Витгенштейна: История десяти минутного спора между двумя великими философами. М., 2004.

Юлииа Н.С. Головоломки проблемы сознания: концепция Дэниела Деннета. М., 2004.

Юлина Н.С. Философский натурализм: о книге Дэниела Деннета «Свобода эволюционирует». М., 2007.

Юм Д. Сочинения в 2т. М., 1996.

Adams E. M. Mental causality // Mind 75 (1966). P. 552-563.

Akins K. Lost the plot? Reconstructing Dennett's multiple drafts theory of consciousness // Mind & Language 11:1 (1996). P. 1-43.

Aldrich V. C, Feigl H, Spatial location and the psycho-physical problem // Philosophy of Science 2: 2 (1935). P. 256-261.

Alter T. The knowledge argument // The Blackwell Companion to Consciousness, ed. by M. Velmans and S. Schneider. Maiden, 2007. P. 396-405.

Antony L. M. I'am a mother, Г worry // Content 6 (1995). P.160-166.

Armstrong D. M. A Materialist Theory of the Mind. Rev. ed. L., 1993.

Armstrong D. M. The Mind-Body Problem: An Opinionated Introduction. Oxford, 1999.

Averill E., Keating B. F. Does interactionism violate a law of classical physics? // Mind 90 (1981). P. 102-107.

Ayer A. J. Language, Truth and Logic. 2 ed. N. Y., 1971 (1946).

Ayer A. J. The Central Questions of Philosophy. N. Y., 1976 (1973).

Ayer A. J. Part of My Life. Oxford, 1978 (1977).

Baars B. A Cognitive Theory of Consciousness. Cambridge, 1993 (1988).

Baars B. In the Theater of Consciousness: The Workspace of the Mind. N. Y., 1997.

Baars B. The global workspace theory of consciousness // The Blackwell Companion to Consciousness, ed. by M. Velmans and S. Schneider. Maiden, 2007, P. 236-246.

Baggini J. *The Pig That Wants to Be Eaten: 100 Experiments for the Armchair Philosopher*. N.Y., 2006 (2005).

Baker L. R. *Metaphysics and mental causation* // *Mental Causation*, ed. by J. Heiland A. Mele. N. Y., 1993. P. 75-95.

Bateson G. *Steps to an Ecology of Mind*. Chicago, 2000.

Bayne T. *Phenomenology and the feeling of doing: Wegner on the conscious will* // *Does Consciousness Cause Behavior?* Ed. by S. Pockett, W. P. Banks and S. Gallagher. Cambridge MA, 2006. P. 168-185.

Beattie J. *An Essay on the Nature and Immutability of Truth in Opposition to Sophistry and Scepticism*. 3 ed. L., 1772.

Beck L. W. *The psychophysical as a pseudo-problem* // *Journal of Philosophy* 37; 21 (1940). P. 561-571.

Beckermann A. *Analytische Einführung in die Philosophie des Geistes*. 2Aufl. B., 2000.

Bennett K. *Mental causation* // *Philosophy Compass* 2: 2 (2007). P. 316-337.

Bennett M. et al. *Neuroscience and Philosophy: Brain, Mind, and Language*. N.Y., 2007.

Blackmore S. *The Meme Machine*. N. Y., 1999.

Blackmore S. *Consciousness: An Introduction*. N. Y., 2004 (2003).

Blackmore S. *Destroying the zombic hunch. Review of Sweet Dreams by Daniel Dennett* // *Nature* 5 (2005) P. 21-23.

Blackmore S. *Conversations on Consciousness: What the Best Minds Think about the Brain, Free Will, and What It Means to Be Human*. N. Y., 2006.

Blackmore S. *Driving beyond consciousness: A test of multiple drafts theory* // *Toward a Science of Consciousness — Tucson 2006*. P. 112-113.

Block N. *Troubles with functionalism* // *Minnesota Studies in the Philosophy of Science* 9 (1978). P. 261-325.

Block N. *On a confusion about a function of consciousness* // *The Nature of Consciousness: Philosophical Debates*, ed. by N. Block, O. Flanagan, and G. Gtzeldere. Cambridge MA, 1997. P. 375-415.

Block N. *The harder problem of consciousness* // *The Journal of Philosophy* 99: 8 (2002). P. 319-425.

Block N. *Review of Daniel Dennett, Consciousness Explained* // *Block N. Consciousness, Function and Representation: Collected Papers, Volume 1*. Cambridge MA, 2007. 129-140.

Bricke J. *Interaction and physiology* // *Mind* 84 (1975). P. 255-259.

Bringsjord S., Noel R. *Real robots and the missing thought-experiment in the Chinese Room dialectic* // *Views into the Chinese Room*, ed. by J. Preston and M. Bishop. N. Y., 2002. P. 144-166.

Broad C D. *Mind and Its Place in Nature*. L., 1925.

Campbell N. *What was Huxley's epiphenomenalism?* // *Biology and Philosophy* 16 (2001). P. 357-375.

Carruthers P. *The Architecture of the Mind*. N. Y., 2006.

Carruthers P. *Higher-order theories of consciousness* // *The Blackwell Companion to Consciousness*, ed. by M. Velinans and S. Schneider. Maiden, 2007. P. 277-286.

Chalmers D. J. *The first-person and third-person views*. P. I. TS 1989. <http://consc.net/notes/first-third.html>.

Chalmers D. J. Consciousness and cognition. TS 1990. <http://consc.net/papers/c-and-c.html>.

Chalmers D. J. Toward a Theory of Consciousness. Ph. D. Thesis. TS 1993.

Chalmers D. J. Absent qualia, fading qualia, dancing qualia // *Conscious Experience*, ed. by T. Metzinger. Paderborn, 1995. P. 309-328.

Chalmers D. The puzzle of conscious experience // *Scientific American* 273 (1995). P. 80-86.

Chalmers D. J. *The Conscious Mind: In Search of a Fundamental Theory*. N.Y., 1997(1996).

Chalmers D. J. Pacing up to the problem of consciousness // *Explaining Consciousness — The "Hard Problem"*, ed. by J. Shear. Cambridge MA, 1997. P. 9-30.

Chalmers D. J. Moving forward on the problem of consciousness // *Explaining Consciousness — The "Hard Problem"*, ed. by J. Shear. Cambridge MA, 1997.. P. 379-422,

Chalmers D. J. Availability: the cognitive basis of experience? // *The Nature of Consciousness: Philosophical Debates*, ed. by N. Block, O. Flanagan, and G. Guzeldere. Cambridge MA, 1997. P 421-424.

Chalmers D. J. On the search for the neural correlate of consciousness // *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates*, ed. by S. R. Hameroff, A. W. Kaszniak, A. C. Scott. Cambridge MA, 1998. P. 219-230.

Chalmers D. J. Materialism and the metaphysics of modality // *Philosophy and Phenomenological Research* 59(1999). P. 473-493.

Chalmers D. J. What is a neural correlate of consciousness // *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, ed. by T. Metzinger. Cambridge MA, 2000. P. 17-40.

Chalmers D. J. Consciousness and its place in nature // *Philosophy of Mind: Classical and Contemporary Readings*, ed. by D. J. Chalmers. N. Y., 2002. P. 247-272.

Chalmers D. J. Ontological anti-realism. TS 2007. <http://consc.net/papers/ontology.pdf>.

Chomsky N. A Review of B. F. Skinner's *Verbal Behavior* // *Language* 35: 1 (1959). P. 26-57.

Chomsky N. Recent contributions to the theory of innate ideas // *Syntheses* 17 (1967). P. 2-11.

Chomsky N. *Reflections on Language*. N.Y., 1975.

Chomsky N. *New Horizons in the Study of Language and Mind*. Cambridge, 2000.

Chomsky N. Turing on the "Imitation game" // *The Turing Test: Verbal Behavior as the Hallmark of Intelligence*, ed. by S. Shieber. Cambridge MA, 2004. P. 317-321.

Churchland P. M. *Matter and Consciousness*. Rev. ed, Cambridge MA, 1988.

Churchland P. M. *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain*. Cambridge MA, 1996 (1995).

Churchland P. M. Densmore and Dennett on virtual machines and consciousness // *Philosophy and Phenomenological Research* 59: 3 (1999). P. 763-767.

Churchland P. M. Knowing qualia: A reply to Jackson (with Postscript 1997) // *There's Something About Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*, ed. by P. Ludow, Y. Nagasawa, and D. Stoljar. Cambridge MA, 2004. P.163-178.

Churchland P. M. Functionalism at forty: a critical retrospective // *Journal of Philosophy* 102:1 (2005). P. 33-50.

Churchland P. S. Can neurobiology teach us anything about consciousness // *The Nature of Consciousness: Philosophical Debates*, ed. by N. Block, O. Flanagan, and G. Guzeldere. Cambridge MA, 1997. P. 127-140.

Churchland P. S. The hornswoggie problem // *Explaining Consciousness — The "Hard Problem"*, ed. by J. Shear. Cambridge MA, 1997. P. 37-44.

Churchland P. S. *Brain-Wise: Studies in Neurophilosophy*. Cambridge MA, 2002.

Clarke C. J. S. The nonlocality of mind // *Explaining Consciousness — The "Hard Problem"*, ed. by J. Shear. Cambridge MA, 1997. P.165-175.

Clarke D. *Descartes's Theory of Mind*. Oxford, 2003.

Cole D. The Chinese Room Argument: Entry in the *Stanford Encyclopedia of Philosophy* 2004 <http://plato.stanford.edu/entries/chinese-room>.

Colvin M., Gazzaniga M. Split-brain cases // *The Blackwell Companion to Consciousness*, ed. by M. Velmans and S. Schneider. Maiden, 2007. P. 181-193.

Corcoran K. J. The trouble with Searle's biological naturalism // *Erkenntnis* 55: 3 (2001). P. 307-324.

Crane T. *Mental causation*. TS 2003.

Crane T. *Intentionalität als Merkmal des Geistigen: Sechs Essays zur Philosophie des Geistes*. Frankfurt am Main, 2007.

Crick F. *The Astonishing Hypothesis: The Scientific Search for the Soul*. N. Y., 1995(1994).

Crick F., Koch C. Towards a neurobiological theory of consciousness // *The Nature of Consciousness: Philosophical Debates*, ed. by N. Block, O. Flanagan, and G. Guzeldere. Cambridge MA, 1997. P. 277-292.

Dahlbom B. (ed.) *Dennett and His Critics: Demystifying Mind*. Maiden, 1993.

Damasio A. *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. Orlando, 2003.

Davidson D. Mental events // *The Nature of Mind*, ed. by D. Rosenthal. P. 247-256.

Dawkins R. *The Selfish Gene*. Oxford, 1999 (1976).

Dehaene S., Naccache L. Toward a cognitive neuroscience of consciousness: basic evidence and a workspace framework // *Cognition* 79(2001). P. 1-37.

Dennett D. C. *Content and Consciousness*. L., 1986 (1969).

Dennett D. C. *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge MA, 1981 (1978).

Dennett D. C. *The Intentional Stance*. Cambridge MA, 1987.

Dennett D.C. *Consciousness Explained*. Boston, 1991.

- Dennett B.C. Darwin's Dangerous Idea: Evolution and the Meanings of Life. L., 1995.
- Dennett D. C. An overview of my work in philosophy. TS 1996.
- Dennett D. C Facing backwards on the problem of consciousness // Explaining Consciousness — The "Hard Problem", ed. by J. Shear. Cambridge MA, 1997. P. 33-36.
- Dennett D. C. Brainchildren: Essays on Designing Minds. L., 1998.
- Dennett D. C. Re-Introducing The Concept of Mind // Ryle G. The Concept of Mind. Chicago, 2002. P. VII-XVII.
- Dennett D. C Freedom Evolves. L., 2004 (2003).
- Dennett D. C. Sweet Dreams: Philosophical Obstacles to a Science of Consciousness. Cambridge MA, 2005.
- Dennett D. C. Interview conducted by Bill Uzgalis in Boston, Massachusetts on December 29, 2004 // Mind and Machines 16:1 (2006). P. 7-19.
- Dennett D. C. Two steps closer to consciousness // Paul Churchland, ed. by B. Keeley. N. Y., 2006. P. 193-209.
- Dennett D. C Intentional systems theory. TS 2007.
- Dennett D. C, Akins K. Multiple drafts model // Scholarpedia 3:4 (2008). P. 4321.
- Dokic J. La philosophie de l'esprit // Precis de philosophie analytique, dir. par P. Engel. Paris, 2000. P. 35-62.
- Dreyfus H. L., Dreyfus S. E. Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer. N. Y., 1988 (1986).
- Dummett M. Origins of Analytical Philosophy. Cambridge MA, 1993.
- Ebert J. J. Unterweisung in den philosophischen und mathematischen Wissenschaften für die oberen Classen der Schulen und Gymnasien. 4 Aufl. Lpz., 1796. S. 593.
- Edelman G. Wider than the Sky: The Phenomenal Gift of Consciousness. L., 2004.
- Edelman G., Tononi G. A Universe of Consciousness: How Matter Becomes Imagination. N. Y., 2000.
- Elitzur A. Consciousness and the incompleteness of the physical explanation of behavior // Journal of Mind and Behavior 10 (1989). P. 1-20.
- Elitzur A. Consciousness makes a difference: Reluctant dualist's confession. TS 2005.
- Elitzur A. Why quantum mind to begin with: A proof for the incompleteness of the physical account of behavior // Quantum Mind — Salzburg 2007. P. 8-9.
- Elton M. Daniel Dennett: Reconciling Science and Our Self-Conception. Cambridge, 2003.
- Faigenbaum G. Conversations with John Searle. LibrosEnRed 2003.
- Feigl G. Logical analysis of the psychophysical problem: A contribution of the new positivism // Philosophy of Science 1: 4 (1934). P. 420-445.
- Feigl G. The "Mental" and the "Physical": The Essay and a Postscript. Minneapolis, 1967.

- Feinberg T.E. *Altered Egos: How the Brain Creates the Self*. N. Y., 2001.
- Feyerabend P. Mental events and the brain // *The Nature of Mind*, ed. by D. Rosenthal. N. Y., 1991. P 266-267.
- Fisher J. Why nothing mental is just in the head // *Nous* 41 (2007). P.318-334.
- Flanagan O. *Consciousness Reconsidered*. Cambridge MA, 1992. Flanagan O. Conscious inessentialism and the epiphenomenalism suspicion // *The Nature of Consciousness: Philosophical Debates*, ed. by N. Block, O. Flanagan, and G. Guzeldere. Cambridge MA, 1997. P. 357-373.
- Flanagan O. *The Problem of the Soul: Two Visions of Mind and How to Reconcile Them*. N. Y., 2002.
- Fodor J. The mind-body problem // *Scientific American* 244 (1981). P. 114-125.
- Fodor J. *The Modularity of Mind*. Cambridge MA, 1983. Fodor J. *A Theory of Content and Other Essays*. Cambridge MA, 1990. Fodor J. *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. Cambridge MA, 2001 (2000).
- Foster J. *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*. L., 1996(1991).
- Gallagher S. Where's the action? Epiphenomenalism and the problem of free will // *Does Consciousness Cause Behavior?* Ed. by S. Pockett, W. P. Banks and S. Gallagher. Cambridge MA, 2006. P. 109-124.
- Gibbons J. Mental causation without downward causation // *Philosophical Review* 115: 1 (2006). P. 79-103.
- Goschke T., Walter H. *Bewusstsein und Willensfreiheit — Philosophische und empirische Annäherungen // Bewusstsein. Phi-losophie, Neurowissenschaften, Ethik*, hrsg. v. C S. Hermann, M. Pauen, J. W. Rieger und S. Schicktanz. Paderborn, 2005. S. 81-119.
- Gray J. *Consciousness: Creeping Up on the Hard Problem*. N. Y., 2006(2004).
- Green J. B., Palmer S. L. (eds) *In Search of the Soul: Four Views of the Mind-Body Problem*. Madison, 2005.
- Guzeldere G. The many faces of consciousness: A field guide // *The Nature of Consciousness: Philosophical Debates*, ed. by N. Block, O. Flanagan, and G. Guzeldere. Cambridge MA, 1997. P. 1-67.
- Hagan S., Hameroff S., Taszynski J. Quantum computation in brain microtubules: Decoherence and biological feasibility // *Physical Review E* 65 061901 (2002).
- Hameroff S., Penrose R. Orchestrated objective reduction of quantum coherence in brain microtubules: The "Orch OR" model for consciousness // *Toward a Science of Consciousness*, ed. by S. Hameroff, A. Kaszniak and A. Scott. Cambridge MA, 1996. P. 507-540.
- Hameroff S. Quantum computation in brain microtubules? The Penrose-Hameroff "Orch OR" model of consciousness // *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* 356: 1743(1998). P. 1869-1896.
- Hameroff S. The good, the bad and the octopus // *Journal of Consciousness Studies* 14: 8 (2007). P. 105-109.

- Hampshire S. The analogy of feeling // *The Nature of Mind*, ed. by D. Rosenthal. N. Y., 1991. P. 98-103.
- Hansen C. M. Between a rock and a hard place: Mental causation and the mind-body problem // *Inquiry* 43: 4 (2000). P. 451-491.
- Hamad S. Grounding symbols in the analog world with neural nets: A hybrid model // *Think* 2 (1993). P. 12-20.
- Hamad S. Minds, machines, and Searle 2: what's right and wrong about the Chinese Room Argument // *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence*, ed. by J. Preston and M. Bishop. N. Y., 2002. P. 294-307.
- Hasker W. *The Emergent Self*. Ithaca, 2001 (1999).
- Hauser L. Searle's Chinese Box: The Chinese Room Argument and Artificial Intelligence. Ph. D. Thesis. 1993. El. ed. <http://members.aol.com/wutsamada/disserta.html>.
- Hawkins J., Blakeslee S. *On Intelligence*. N. Y., 2004.
- Heidelberger M. Wie das Leib-Seele Problem in den logischen Empirismus kam // *E-Journal Philosophie der Psychologie*. 2 (2005). <http://www.jp.philo.at/texte/HeidelbergerMI.pdf>.
- Heil J. Mental causation // *The Blackwell Guide to Philosophy of Mind*, ed. by S. P. Stich and T. A. Warfield. Maiden, 2003. P. 214-234.
- Herrmann C. S., Pauen M., Mm B. K., Busch N.A., Rieger J. W. Eine neue Interpretation von Libets Experimenten aus der Analyse einer Wahlreaktionsaufgabe // *Bewusstsein. Philosophie, Neurowissenschaften, Ethik*, hrsg. v. C. S. Hermann, M. Pauen, J. W. Rieger und S. Schicktanz. Paderborn, 2005. S. 120-134.
- Hissmann M. *Psychologische Versuche. Ein Beitrag zur esoterischen Logik*. Frankfurt am Main., 1777.
- Hodgson D. *The Mind Matters: Consciousness and Choice in a Quantum World*. N. Y., 1991.
- Hodgson D. The easy problems ain't so easy // *Explaining Consciousness — The "Hard Problem"*, ed. by J. Shear. Cambridge MA, 1997. P. 125-131.
- Hofstadter D., Dennett D. *The Mind's I: Fantasies and Reflections on Self and Soul*. N.Y., 2000 (1981).
- Hofstadter D. I am a Strange Loop. N. Y., 2007.
- Hogan T. Supervenient qualia // *The Philosophical Review* 96: 4 (1987). P. 491-520.
- Horowitz A. Is there a problem in physicalist epiphenomenalism? // *Philosophy and Phenomenological Research* 59: 2 (1999). P. 421-434.
- Humphrey N. How to solve the mind-body problem // *Journal of Consciousness Studies* 7 (2000). P. 5-20.
- Humphrey N. *Seeing Red: A Study in Consciousness*. Cambridge MA, 2006.
- Huxley T. On the hypothesis that animals are automata, and its history // *Fortnightly Review* (November 1874). P. 555-580.
- Jackendoff R. *Consciousness and the Computational Mind*. Cambridge MA, 1987.
- Jackson F. Epiphenomenal qualia // *There's Something About Mary*, ed. by P. Ludow, Y. Nagasawa, and D. Stoljar. Cambridge MA, 2004. P. 39-50.

- Jackson F. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. N. Y., 1998.
- Jackson F, *Mind and illusion // There's Something About Mary*, ed. by P. Ludow, Y. Nagasawa, and D. Stoljar. Cambridge MA, 2004. P. 421-442.
- James W. *Are we automata? // Mind* 4 (1879). P. 1-22.
- Kant I. *Gesammelte Schriften. Akademie-Ausgabe. B.*, 1900ff.
- Kenny A. *The Metaphysics of Mind*. Oxford, 1989.
- Kim J. *Epiphenomenal and supervenient causation // The Nature of Mind*, ed. by D. Rosenthal. N. Y., 1991. P. 257-285.
- Kim J. "Strong" and "global" supervenience revisited // *Supervenience and Mind: Selected Philosophical Essays*, ed. by J. Kim. Cambridge, 1993. P. 79-91.
- Kim J. *Mental causation in Searle's "biological naturalism" // Philosophy and Phenomenological Research* 55; 1 (1995). P. 189-194.
- Kim J. *The mind-body problem: Taking stock after forty years // Philosophical Perspectives* 11 (1997). P. 185-207.
- Kim J. *Mind in a Physical World*. Cambridge MA, 1998.
- Kim J. *Physicalism, or Something Near Enough*. Princeton, 2005.
- Kim J. *Philosophy of Mind*. Cambridge MA, 2006.
- Kim J. *The Causal efficacy of consciousness // The Blackwell Companion to Consciousness*, ed. by M. Velmans and S. Schneider. Maiden, 2007. P. 406-417.
- King P. *Why isn't the mind-body problem medieval // Forming the Mind: Essays on the Internal Senses and the Mind/Body Problem from Avicenna to the Medical Enlightenment*, ed. by H. Lagerlund. Dordrecht, 2007. P. 187-206.
- Kirk R. *Sentience and behaviour // Mind* 83 (1974). P. 43-60.
- Kirk R. "The best set of tools"? Dennett's metaphors and the mind-body problem // *The Philosophical Quarterly* 43:172 (1993). P. 335-343.
- Kirk R. *Zombies and Consciousness*. N.Y., 2005.
- Kitcher P. *Kant's Transcendental Psychology*. N. Y., 1990.
- Koch C *The Quest for Consciousness: A Neurobiological Approach*. Englewood, 2004.
- Koksvik O. *In Defence of Interactionism*. Master thesis. TS 2006.
- Kriegel U., Williford K. (eds) *Self-Representational Approaches to Consciousness*. Cambridge MA, 2006.
- Kripke S. *Naming and Necessity*. Cambridge MA, 1980.
- Kripke S. *Wittgenstein on Rules and Private Language*. Cambridge MA, 1982.
- Kuczynski J.-M. M. *A quasi-materialist, quasi-dualist solution to the mind-body problem // Criterion* 109 (2004). P. 81-135.
- Lacks J. *Epiphenomenalism and the notion of cause // The Journal of Philosophy* 60 (1963). P. 141-146.
- Leopold D. A., Logothetis N. K. *Activity changes in early visual cortex reflects monkeys' percepts during binocular rivalry // Nature* 379 (1996). P. 549-553.
- Levine J. *Purple Haze: The Puzzle of Consciousness*. N. Y., 2004 (2001).
- Levine J. *Anti-materialist arguments and influential replies // The Blackwell Companion to Consciousness*, ed. by M. Velmans and S. Schneider. Maiden, 2007. P. 371-380.

- Lewis C. I. *Mind and the World-Order: Outline of a Theory of Knowledge*. N. Y., 1929.
- Lewis D. An argument for the identity theory // *Journal of Philosophy* 63: 1 (1966). P. 17-25.
- Lewis D. *Philosophical Papers*. V. II. N. Y., 1986.
- Lewis D. Mad pain and Martian pain // *The Nature of Mind*, ed. by D. Rosenthal. N. Y., 1991. P. 229-234.
- Lewis D. What experience teaches // *There's Something About Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*, ed. by P. Ludow, Y. Nagasawa, and D. Stoljar. Cambridge MA, 2004. P. 77-103.
- Libet B. Solutions to the hard problem of consciousness // *Explaining Consciousness — The "Hard Problem"*, ed. by J. Shear. Cambridge MA, 1997. P. 301-303.
- Libet B. *Mind Time: The Temporal Factor in Consciousness*. Cambridge MA, 2004.
- Loar B. Phenomenal states // *The Nature of Consciousness: Philosophical Debates*, ed. by N. Block, O. Flanagan, and G. Guzeldere. Cambridge MA, 1997. P. 597-616.
- Lowe E. J. There are no easy problems of consciousness // *Explaining Consciousness — The "Hard Problem"*, ed. by J. Shear. Cambridge MA, 1997. P. 117-123.
- Lowe E. J. Causal closure principles and emergentism // *Philosophy* 75: 4 (2000). P. 571-585.
- Malcolm N. Knowledge of other minds // *The Nature of Mind*, ed. by D. Rosenthal. N. Y., 1991. P. 92-97.
- Marcus E. Mental causation: Unnaturalized but not unnatural // *Philosophy and Phenomenological Research* 61:1 (2001). P. 57-83.
- Marras A. Methodological and ontological aspects of the mental causation problem // *Physicalism and Mental Causation: The Metaphysics of Mind and Action*, ed. by S. Walter and H.-D. Heckmann. Exeter, 2003. P. 243-264.
- Marsoobian A. T., Ryder J. (eds) *The Blackwell Guide to American Philosophy*. Maiden, 2004.
- McGinn C. *The Problem of Consciousness: Essays Towards a Resolution*. Oxford, 1993(1991).
- McGinn C. *Problems in Philosophy: The Limits of Inquiry*. Maiden, 1993.
- McGinn C. Consciousness evaded: Comments on Dennett // *Philosophical Perspectives* 9 (1995). P. 241-249.
- McGinn C. *Character of Mind: An Introduction to the Philosophy of Mind*. New ed. N. Y., 1996.
- McGinn C. Can we solve the mind-body problem? // *The Nature of Consciousness: Philosophical Debates*, ed. by N. Block, O. Flanagan, and G. Guzeldere. Cambridge MA, 1997. P. 529-542.
- McGinn C. *The Mysterious Flame: Conscious Minds in a Material World*. N. Y., 1999.
- McGinn C. *The Making of a Philosopher: My Journey through Twentieth-Century Philosophy*. N. Y., 2003 (2002).
- McGinn C. *Consciousness and Its Objects*. N.Y., 2004.

McGinn C. Hard questions // Strawson G. et al. *Consciousness and Its Place in Nature: Does Physicalism Entail Panpsychism?* Charlottesville, 2006. P. 90-99.

McKinsey M. Refutation of qualia-physicalism // *Situating Semantics: Essays on the Philosophy of John Perry*, ed. by M. O'Rourke and Corey Washington. Cambridge MA, 2007. P. 469-498.

McLaughlin B. P. Type epiphenomenalism, type dualism, and the causal priority of the physical // *Philosophical Perspectives* 3 (1989). P. 109-135.

Meixner U. New perspectives for a dualistic conception of mental causation // *Journal of Consciousness Studies* 15: 1 (2008). P. 17-38.

Melloni L. et al. Synchronization of neural activity across cortical areas correlates with conscious perception // *The Journal of Neuroscience* 27: 11 (2007). P. 2858-2865.

Menzies P. Counterfactual theories of causation: Entry in the *Stanford Encyclopedia of Philosophy*, 2001. <http://plato.Stanford.edu/entries/causation-counterfactual>.

Menzies P. The causal efficacy of mental states // *Physicalism and Mental Causation: The Metaphysics of Mind and Action*, ed. by S. Walter and H.-D. Heckmann. Exeter, 2003. P. 195-223.

Meyer M. The Present status of the problem of the relation between mind and body // *The Journal of Philosophy, Psychology and Scientific Methods* 9: 14(1912). P. 365-371.

Mills E. O. Giving up on the hard problem of consciousness // *Explaining Consciousness — The "Hard Problem"*, ed. by J. Shear. Cambridge MA, 1997. P. 109-115.

Mohrhoff U. Interactionism, energy conservation, and the violation of physical laws // *Physics Essays* 10: 4 (1997). P. 651-665.

Monk R. *Ludwig Wittgenstein: The Duty of Genius*. N. Y., 1991 (1990).

Montero B. Varieties of causal closure // *Physicalism and Mental Causation: The Metaphysics of Mind and Action*, ed. by S. Walter and H.-D. Heckmann. Exeter, 2003. P. 173-187.

Montero B. What does the conservation of energy have to do with physicalism? // *Dialectica* 60: 4 (2006). P. 383-396.

Moore G. E. *Some Main Problems of Philosophy*. N. Y., 1962 (1953).

Moural J. The Chinese Room Argument // *John Searle*, ed. by B. Smith. Cambridge, 2003. P. 214-260.

Muijnck W. *Dependencies, Connections, and Other Relations: A Theory of Mental Causation*. Dordrecht, 2003.

Muijnck W. Two types of mental causation // *Philosophical Explorations*, 7:1 (2004). P. 21-35.

Nagel T. Brain bisection and the unity of consciousness // *Synthese* 22(1971). P. 396-413.

Nagel T. What is it like to be a bat? // *Philosophical Review* 83: 4 (1974). P. 435-450.

Nagel T. *The View From Nowhere*. N. Y., 1986.

Nagel T. *Other Minds: Critical Essays*. 1969-1994. N. Y., 1995.

Papineau D. *Thinking about Consciousness*. N. Y., 2004 (2002).

Parfit D. Personal identity // *Personal Identity*, ed. by J. Perry. Berkeley, 1975. P. 199-223.

- Parfit D. *Reasons and Persons*. N. Y., 1986 (1984).
- Passingham R. E., Lau H. C, Free choice and the human brain // *Does Consciousness Cause Behavior?* Ed. by S. Pocketi, W. P. Banks and S. Gallagher. Cambridge MA, 2006. P. 53-72.
- Patten M., Staudacher A., Walter S. Epiphenomenalism: Dead end or way out? // *Journal of Consciousness Studies* 13: 1-2 (2006). P. 7-19.
- Pauen M. *Was ist der Mensch? Die Entdeckung der Natur des Geistes*. Mtinchen, 2007.
- Pinker S. *How the Mind Works*. L., 1999(1997).
- Pinker S. *The Blank Slate: The Modern Denial of Human Nature*. N.Y.,2002.
- Pinker S. So how does the mind work? // *Mind & Language* 20:1 (2005). P. 1-24.
- Pinker S- *The Stuff of Thought: Language as a Window into Human Nature*. N. Y., 2007.
- Place U. T. Is consciousness a brain process? // *British Journal of Psychology* 47 (1956). P. 44-50.
- Place U. T. Token— versus type-identity physicalism // *Antropology and Philosophy* 3: 2 (1999). P. 21-31.
- Place U. T. *Identity theories*. TS 1999.
- Place U. T. *Identifying the Mind: Selected Papers of U.T. Place*. N. Y., 2004.
- Polger T. Rethinking the evolution of consciousness // *The Blackwell Companion to Consciousness*, ed. by M, Velmans and S- Schneider. Maiden, 2007. P. 72-86.
- Popper K., Eccles J. *The Self and Its Brain: An Agrument for Interactionism*. B., 1977.
- Pratt J. B. The present status of the mind-body problem // *The Philosophical Review* 45: 2 (1936). P. 144-166.
- Preston A. Prolegomena to any future history of analytic philosophy // *Metaphilosophy* 35: 4 (2004).
- Preston J., Bishop M. (eds) *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence*. N. Y., 2002.
- Priestley J. *Disquisitions relating to Matter and Spirit*. L., 1777.
- Prinz J- The intermediate level theory of consciousness // *The Blackwell Companion to Consciousness*, ed. by M. Velmans and S. Schneider. Maiden, 2007. P. 247-260.
- Putnam H. Minds and machines // *Dimensions of Mind*, ed. by S. Hook. N. Y., 1960. P. 148-180.
- Putnam H. Psychological predicates // *Art, Mind, and Religion*, ed. by W. Capitan and D. Merrill. Pittsburgh, 1967. P. 37-48.
- Putnam H. *Words & Life*. Cambridge MA, 1994.
- Quine W. V. *The Ways of Paradox*. N. Y., 1966.
- Quine W. V. *Theories and Things*. Cambridge MA, 1981.
- Quine W. V. *Pursuit of Truth*. Cambridge MA, 1990.
- Quine W, V. States of mind // *The Nature of Mind*, ed. by D. Rosenthal. N. Y., 1991. P. 287-288.
- Ramachandran V. S., Hirstein W. Three laws of qualia: What neurology tells us about the biological functions of consciousness, qualia and the self // *Journal of Consciousness Studies* 4: 4-5 (1998). P. 429-457.

Ramachandran V. S. Mirror neurons and imitation learning as the driving force behind "the great leap forward" in human evolution // *Edge* (2000).

Ramachandran V. S. *A Brief Tour of Human Consciousness: From Impostor Poodles to Purple Numbers*. N. Y., 2004.

Reid T. *The Works of Thomas Reid*. 7 ed. Edinburgh, 1872.

Richardson- R. C The "scandal" of Cartesian interactionism // *Mind* 91(1982). P. 20-37.

Rieger J. W., Schichtanz S. "Wenn Du denkst, dass ich denke, dass Du denkst..." — Anmerkungen zur interdisziplinären Auseinandersetzung über das Bewusstsein // *Bewusstseiu. Philosophie, Neurowissenschaften, Ethik*, hrsg. v. C S. Hermann, M. Pauen, J. W. Rieger und S. Schicktanz. Paderborn, 2005. S. 22-52.

Rivas T., Dongen H. v. Exit epiphenomenalism: The demolition of a refuge // *JNL RMI* 2:1 (2003).

Robb D., Hell J. Mental causation: Entry in the *Stanford Encyclopedia of Philosophy*, 2003. <http://plato.stanford.edu/entries/mental-causation>.

Robinson D. N. *Consciousness and Mental Life*. N. Y., 2008.

Robinson H. Dennett on the Knowledge argument // *There's Something About Mary*, ed. by P. Ludow, Y. Nagasawa, and D. Stoljar. Cambridge MA, 2004. P. 69-73.

Robinson W. S. The hardness of the hard problem // *Explaining Consciousness — The "Hard Problem"*, ed. by J. Shear. Cambridge MA, 1997. P. 149-161.

Robinson W. S. Epiphenomenalism: Entry in the *Stanford Encyclopedia of Philosophy*, 2007. <http://plato.stanford.edu/entries/epiphenomenalism>.

Robinson W. S. Evolution and epiphenomenalism // *Journal of Consciousness Studies* 14: 11 (2007). P. 27-42.

Rosenberg G.H. Rethinking nature: A hard problem within the hard problem // *Explaining Consciousness — The "Hard Problem"*, ed. by J. Shear. Cambridge MA, 1997. P. 287-300.

Rosenberg G. H. *A Place for Consciousness: Probing the Deep Structure of the Natural World*. N. Y., 2004.

Rosenblum B., Kuttner F. *Quantum Enigma: Physics Encounters Consciousness*. N. Y., 2006.

Rosenthal D. Two concepts of consciousness // *The Nature of Mind*, ed. by D. Rosenthal. N. Y., 1991. P. 462-477.

Rosenthal D. M. A Theory of consciousness // *The Nature of Consciousness: Philosophical Debates*, ed. by N. Block, O. Flanagan, and G. Guzeldere. Cambridge MA, 1997. P. 729-753.

Rosenthal D. M. *Consciousness and Mind*. N. Y., 2005.

Ross J. A. Hitting on consciousness: Honderich versus McGinn // *Journal of Consciousness Studies* 15: 1 (2008). P. 109-128.

Rowlands M. Mysterianism // *The Blackwell Companion to Consciousness*, ed. by M. Velmans and S. Schneider. Maiden, 2007. P. 335-345.

Russell B. *The Analysis of Matter*. New ed. N. Y., 1954.

Russell B. Analogy // *The Nature of Mind*, ed. by D. Rosenthal. N. Y., 1991. P. 89-91.

Ryle G. *The Concept of Mind*. Chicago, 2002 (1949).

- Schneider S. Daniel Dennett on the nature of consciousness // *The Blackwell Companion to Consciousness*, ed. by M. Velmans and S. Schneider. Maiden, 2007. P. 313-324.
- Schutt H.-P. *Substanzen, Subjekte und Personen: Eine Studie zum Cartesischen Dualismus*. Heidelberg, 1990.
- Searle J. R. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge, 1969.
- Searle J. R. *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge, 1979.
- Searle J. R. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge, 1983.
- Searle J. R. *Minds, Brains and Science*. Cambridge MA, 1984.
- Searle J. R. *Minds, Brains, and Programs // The Nature of Mind*, ed. by D. Rosenthal. N. Y., 1991. P. 509-519.
- Searle J. R. Response: The mind-body problem // *John Searle and His Critics*, ed. by E. Lepore and R. Van Gulick. Cambridge MA, 1991. P. 141-146.
- Searle J. R. *The Rediscovery of the Mind*. Cambridge MA, 1994 (1992).
- Searle J. R. Consciousness, the brain and the connection principle: A reply // *Philosophy and Phenomenological Research* 55: 1 (1995). P. 217-232.
- Searle J. R. *The Mystery of Consciousness*. L., 1998(1997).
- Searle J. R. *Mind, Language, and Society: Philosophy in the Real World*. N.Y., 1999(1998).
- Searle J. R. The future of philosophy // *Philosophical Transactions: Biological Sciences* 354: 1392 (1999). P. 2069-2080.
- Searle J. R. Consciousness // *Annual Review of Neuroscience* 23 (2000). P. 557-578.
- Searle J. R. Why I am not a property dualist // *Journal of Consciousness Studies*, 9: 12 (2002) P. 57-64.
- Searle J. R. *Consciousness and Language*. Cambridge, 2002.
- Searle J. R. *Mind: A Brief Introduction*. Oxford, 2004.
- Searle J. R. *Freedom and Neurobiology: Reflections on Free Will, Language, and Political Power*. N. Y., 2007.
- Searle J. R. Biological naturalism // *The Blackwell Companion to Consciousness*, ed. by M. Velmans and S. Schneider. Maiden, 2007. P.325-334.
- Sellars W. *Empiricism and the Philosophy of Mind*. Cambridge MA, 1997.
- Shaffer J. Mental events and the brain // *The Nature of Mind*, ed. by D. Rosenthal. N. Y., 1991. P. 177-180.
- Shoemaker S. Functionalism and qualia // *The Nature of Mind*, ed. by D. Rosenthal. N.Y., 1991. P. 395-407.
- Shoemaker S. The mind-body problem // *Minds and Bodies: An Introduction with Readings*, ed. by R. Wilkinson. N. Y., 2000. P. 194-197.
- Singer W. Large-scale temporal coordination of cortical activity as a prerequisite for conscious experience // *The Blackwell Companion to Consciousness*, ed. by M. Velmans and S. Schneider. Maiden, 2007. P. 605-615.
- Skinner B. F. *Verbal Behavior*. Acton, 1957.
- Smart J. J. C. Sensations and brain processes // *The Nature of Mind*. ed. by D. Rosenthal. P. 169-176.

- Smart J. J. C. *Postscript* (1995) // *Philosophy of Mind: Contemporary Readings*, ed. by T. O'Connor and D. Robb. N. Y., 2003.
- Smart J. J. C. The Identity theory of mind: Entry in the *Stanford Encyclopedia of Philosophy*, 2007. <http://plato.stanford.edu/entries/mind-identity/#His>.
- Soames S. *Philosophical Analysis in the Twentieth Century*. V. 1-2. Princeton, 2005 (2003).
- Sorensen R.A. *Thought Experiments*. N. Y., 1992.
- Sosa E. Mind-body interaction and supervenient causation // *Midwest Studies in Philosophy* 9 (1984). P. 271-281.
- Sparacio M. Mental realism: Rejecting the causal closure thesis and expanding our physical ontology // *PCID* 2.3.8 (2003).
- Spat P. A pill against epiphenomenalism // *Abstracta* 2: 2 (2006). P. 172-179.
- Sperry R. W. Hemisphere disconnection and unity in conscious awareness // *American Psychologist* 23 (1968). P. 723-733
- Sperry R. In defense of mentalism and emergent interaction // *The Journal of Mind and Behavior* 12: 2 (1991). P. 221-245.
- Standage T. *Der Turke: Die Geschichte des ersten Schachautomaten und seiner abenteuerlichen Reise um die Welt*. Berlin, 2005.
- Staudacher A. Epistemological challenges to qualia-epiphenomenalism // *Journal of Consciousness Studies* 13: 1-2 (2006). P. 153-175.
- Stoerig P., Cowey A. Blindsight in man and monkey. *Brain* 120 (1997). P. 535-559.
- Stoljar D. *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness*. N.Y., 2006.
- Strawson G. *Mental Reality*. Cambridge MA, 1994.
- Straivson G. Realistic monism: Why physicalism entails panpsychism // Strawson G. et al. *Consciousness and Its Place in Nature: Does Physicalism Entail Panpsychism?* Charlottesville, 2006. P. 3-31.
- Strawson P. F. *Individuals: An Essay in Descriptive Metaphysics*. L., 1964(1959).
- Strawson P. F. *The Bounds of Sense: An Essay on Kant's Critique of Pure Reason*. L., 1966.
- Strawson P. F. *Analysis and Metaphysics: An Introduction to Philosophy*. N. Y., 1992.
- Stroll A. *Twentieth-Century Analytic Philosophy*. N. Y., 2000.
- Sturma D. *PhilosophiedesGeistes*. Leipzig, 2005.
- Swinburne R. *The Evolution of the Soul*. Rev. ed. Oxford, 1997.
- Tegmark M. The importance of quantum decoherence in brain processes // *Physical Review E* 61 (2000). P. 4194-4206.
- Tetens J. N. *Philosophische Versuche ber die menschliche Natur und ihre Entwicklung*. Bd. 1-2 Lpz., 1777.
- Turing A. Computing machinery and intelligence // *Mind* 59 (1950). P.433-460.
- Tuszynski J. (ed.) *The Emerging Physics of Consciousness*. B., 2006.
- Tye M. The function of consciousness // *Nous* 30: 3 (1996). P. 287-305.
- Tye M. Knowing what it is like: The ability hypothesis and the Knowledge argument//*There's Something About Mary*, ed. by P. Ludow, Y. Nagasawa, and D. Stoljar. Cambridge MA, 2004. P. 143-160.

- Van Gulik R. Functionalism and qualia // *The Blackwell Companion to Consciousness*, ed. by M. Velmans and S. Schneider. Malden, 2007. P. 381-395.
- Vasilyev V. Brain and consciousness: Exits from the labyrinth // *Social Sciences* 37: 2 (2006). P. 51-66.
- Vasilyev V. Mental causation, common sense and quantum mechanics // *Quantum Mind — Salzburg* 2007. P. 77-79.
- Vasilyev V. Interactionism is compatible with causal closure // *Toward a Science of Consciousness — Tucson* 2008. P. 55.
- Velmans M. *Understanding Consciousness*. L., 2000.
- Velmans M. *How Could Conscious Experiences Affect Brains*. Charlottesville, 2002.
- Velmans M. Dualism, reductionism, and reflexive monism // *The Blackwell Companion to Consciousness*, ed. by M. Velmans and S. Schneider. Maiden, 2007. P. 346-358.
- Volkov D. B. The moving target: Multiple drafts or fame in brain? *TS* 2007.
- Voorhees B. Dennett and the deep blue sea // *Journal of Consciousness Studies* 7: 3 (2000). P. 53-69.
- Walter S. Epiphenomenalism: Entry in the Internet Encyclopedia of Philosophy, 2007. <http://www.iep.utm.edu/e/epiplieno.htm#H2>.
- Walter S. Ist der Epiphanomenalismus absurd? Ein frischer Blick auf eine tot geglaubte Position. *TS* 2007.
- Wassermann G. D. Reply to Popper's attack on epiphenomenalism // *Mind* 88 (1979). P. 572-575.
- Wegner D.M. *The Illusion of Conscious Will*. Cambridge MA, 2003 (2002).
- Weiskrantz L. The case of blindsight // *The Blackwell Companion to Consciousness*, ed. by M. Velmans and S. Schneider. Maiden, 2007. P.175-180.
- Whitehead Ch. Everything I believe might be a delusion. Whoa! *Tucson* 2004: Ten years on, and are we any nearer to a science of consciousness? // *Journal of Consciousness Studies* 11: 12 (2004). P.68-88.
- Whyte L. L. *The Unconscious before Freud*. N. Y, 1960.
- Williamson T. *The Philosophy of Philosophy*. Maiden MA, 2007.
- Wolff Chr. *Vernunfftige Gedanketi von Gott, der Welt und der Seele des Menschen, auch alien Dingen uberhaupt*. 3 Aufl. Halle, 1725; *Anderer Theil*. Andere Aufl. Frankfurt am Main, 1727.
- Wright J. P., Potter P. (eds) *Psyche and Soma: Physicians and Metaphysicians on the Mind-Body Problem from Antiquity to Enlightenment*. N. Y., 2000.
- Wright W, Explanation and the hard problem // *Philosophical Studies* 132 (2007). P. 301-330.
- Yablo S. Mental causation // *The Philosophical Review* 101:2 (1992). P. 245-280.
- Zawidzki T. *Dennett*. Glasgow, 2007.
- Zimmiermann F. J. *Philosophisch-medicinisches Worterbuch*. Wien, 1803.
- Zizek S. *The Parallax View*. Cambridge MA, 2006.