

«Детектив — это расследование тайны, замешанной на преступлении, насилии, крови. Как насчет аналитического детектива, завязанного на самых интригующих тайнах нашей души?» — профессор К. Анохин

Свобода ВОЛИ

Иллюзия или возможность



Дмитрий Волков



Центр исследования
сознания

ПРИ ФИЛОСОФСКОМ ФАКУЛЬТЕТЕ МГУ

0+

Дмитрий Волков
Свобода воли. Иллюзия
ИЛИ ВОЗМОЖНОСТЬ

*http://www.litres.ru/pages/biblio_book/?art=56036454
SelfPub; 2020*

Аннотация

Многие открытия в нейронауке и социальной психологии ставят под сомнение наличие свободы воли. Они показывают, что поступки людей определяются социальным контекстом и бессознательными состояниями мозга. Но не следует ли из этого, что человек не может принимать решения сам? А если так, если свобода воли – иллюзия, возможно ли считать других людей ответственными за поведение? В книге автор защищает современную компатибилистскую теорию – позицию, согласно которой свобода и ответственность совместимы с современными научными представлениями об устройстве мира. Кроме того, в ней предложены решения и других философских головоломок, связанных со свободой: проблемы тождества личности и ментальной каузальности. Эта книга – результат многолетних исследований автора. Дмитрий Борисович Волков – доктор философских наук, содиректор Московского центра исследования сознания при философском факультете МГУ,

автор книги «Бостонский зомби: Д. Деннет и его теория сознания» (2012).

Содержание

Предисловие	15
Введение	20
Глава 1. Проблема ментальной каузальности и возможность действий	49
I. Аргументы в пользу ментальной каузальности	49
1. Связь проблем ментальной каузальности и свободы воли	49
2. Корреляция ментальных состояний и действий	51
3. Два типа возможных связей между ментальными состояниями и действиями	56
4. Доводы против каузальной связи действий и ментальных состояний	59
А. Трудности с доказательством психофизических законов	59
В. Аргумент аномальности ментального	62
С. Аргумент логической связи	63
5. Доводы Д. Дэвидсона в пользу каузальной связи	65
II. Проблема ментальной каузальности в современной аналитической философии	74
1. Аргумент исключения ДЖ. Кима против	74

ментальной каузальности	
2. Аргумент супервентности против каузального влияния одних ментальных состояний на другие	80
III. Решение проблемы ментальной каузальности в локальном интеракционизме	85
1. Ключевые положения локального интеракционизма	85
2. Аргумент каузальных траекторий в пользу ментальной каузальности и против локальной супервентности	87
А. Локальная или глобальная супервентность?	93
В. Разъяснение двусмысленности в аргументе	98
3. Возражения на аргумент каузальных траекторий	101
А. Аргумент-двойник	101
В. Память не является качественным состоянием	107
С. Решение парадокса сближающихся в ментальных состояниях двойников	110
IV. Решение проблемы ментальной каузальности в биологическом натурализме Дж. Сёрла	115
1. Ключевые положения биологического	115

натурализма Дж. Сёрла	
2. Иллюстрации к каузальным связям в биологическом натурализме	121
3. Внутренние противоречия теории Дж. Сёрла	127
4. Недостатки проведенных аналогий и иллюстраций	129
V. Решение проблемы ментальной каузальности в телеофункционализме Д. Деннета	132
1. Ключевые положения телеофункционализма Деннета	132
2. «История двух черных ящиков» как иллюстрация каузальной эффективности высокоуровневых свойств	141
3. Интерпретация мысленного эксперимента	146
4. Критика телеофункционализма и возможные ответы	151
А. Невыполнимость условий мысленного эксперимента	151
В. Машина не обладает семантическими свойствами	153
С. Высокоуровневые свойства в принципе не обладают каузальной эффективностью	160

D. Функциональные свойства не обладают каузальной эффективностью	163
VI. Решение проблемы ментальной каузальности	171
1. Аргумент обобщения Н. Блока против аргумента исключения	171
2. Контрпример к аргументу исключения. Библиотека первых изданий	181
3. Что не так с аргументом исключения Дж. Кима?	188
VII. Итоги первой главы	192
Глава 2. Свобода воли и детерминизм	194
I. Проблема соотношения свободы, детерминизма и индетерминизма	194
A. Ответственность как индикатор свободы	196
B. Действия без принуждения не обязательно являются свободными и ответственными	202
C. Два принципа свободы воли	209
1. Свобода воли и детерминизм	213
A. Каузальный детерминизм	213
B. Два аргумента против свободы в условиях детерминизма	215
C. Каузальный детерминизм и квантовая физика	219
2. Свобода воли и индетерминизм.	227

Аргумент случайности	
3. Дилемма свободы воли: соотношение с детерминизмом и индетерминизмом	232
4. Альтернативные позиции по вопросу свободы воли конца XX – начала XXI В	234
II. Решение проблемы свободы воли в либертарианстве Р. Кейна	237
1. Этапы либертарианской кампании	237
2. Свобода воли и детерминизм несовместимы	240
А. Принцип альтернативных возможностей	241
В. Принцип автономии и «ответственность в конечном счете»	248
С. Соотношение двух принципов свободы в теории кейна	252
3. Свобода воли и индетерминизм совместимы	257
А. Контрпримеры к классическому аргументу случайности	257
В. Индетерминизм и расширенный контроль	260
4. Возражения к теории Кейна	266
А. Практическая значимость и возможность волеустанавливающих действий	268

В. Отсутствие ответственности в конечном счете в случае ВУД	270
III. Свобода воли и моральная ответственность в жестком инкомпатибилизме	277
Д. Перебума	
1. Свобода воли и индетерминизм несовместимы	279
А. Индетерминизм не расширяет контроль	279
В. Аргумент исчезающего агента	292
2. Свобода воли и детерминизм несовместимы: аргумент манипуляций	297
IV. Обоснование суперкомпатибилизма	307
1. Возражение на Аргумент манипуляций	307
2. Возражение на аргумент исчезающего агента	323
А. А. Мили	324
В. К. Франклин: Усовершенствованный вариант событийно-каузального либертарианства позволяет устанавливать выбор	328
С. Личность как пердурантная сущность устанавливает выбор	331
V. Итоги второй главы	341
Глава 3. Проблема тождества личности и моральная ответственность	344

I. Событийный каузализм как угроза тождеству личности	344
1. Связь вопроса о тождестве личности с проблемами моральной ответственности и свободы	344
2. Событийный каузализм	347
3. Проблема единства и тождества личности	351
II. Современные классические теории тождества личности	355
1. Психологический подход	357
А. Дж. Локк: память как критерий тождества личности	358
В. Возражения к теории памяти Дж. Локка	361
С. Современные вариации психологического подхода	364
D. Аргумент редупликации и Аргумент телепортации	368
Е. Решение проблемы редупликации в современных психологических теориях	372
2. Биологический подход Э. Олсона	375
А. Аргументы Э. Олсона в пользу биологического подхода	376
В. Сравнение биологического, физиологического и психологического	381

подходов	
С. Критика биологического подхода	383
3. Субстанциональный подход и его сторонники	385
А. Субстанциональный поход Р. Суинбёрна	387
В. Недостатки эмпирицистских теорий	389
С. Особенности субстанционального подхода суинбёрна	392
Д. Критика субстанционального подхода с помощью верификационизма	395
III. Критика классических подходов с помощью Аргумента пяти трансплантаций	402
А. Трансплантация души	402
В. Трансплантация мозга	404
С. Трансплантация психики	405
Д. Трансплантация психики с умножением личности	406
Е. Трансплантация психики с делением личности (Fission case)	407
IV. Теория личности Д. Деннета	410
1. Мысленный эксперимент «Где Я?»	410
Конфигурация 1. Гамлет (тело) и Йорик (мозг)	412
Конфигурация 2. Йорик (старый мозг) и Фортинбрас (новое тело)	417

Конфигурация 3. Йорик (старый мозг) и Фортинбрас / Губерт (новый мозг) и Фортинбрас, Йорик = Губерт	418
Конфигурация 4. Йорик и Фортинбрас / Губерт и Фортинбрас, Йорик \neq Губерт; New fission case	421
2. Личность как результат эволюции	425
А. Появление границ, предпочтений и агентов	425
В. Личность как расширенный фенотип	428
3. Критерии личности	432
А. Коллизия моральной и юридической ответственности в случаях психопатии	439
В. Психопатия как расстройство личности	441
С. Моральное чувство инопланетян	451
D. Община джайнов и ахимса	453
Е. Психопаты и закоренелые преступники	455
V. Нарративный подход к вопросу тождества личности	460
1. Конструкция характеризующего биографического нарратива	463
2. Преимущества нарративного подхода	470
3. Возражения к нарративному подходу	473
Возражение 1. Проблема двойников	473

Возражение 2. Личность предшествует нарративу	475
Возражение 3. Проблема кого-то еще	477
Возражение 4. Эпизодические личности	479
VI. Личность как центр нарративной гравитации	487
1. Полезные аналогии, проясняющие, что такое личность	487
2. Эмпирические доводы в пользу концепции Центра нарративной гравитации	493
А. Диссоциативное расстройство личности	493
В. Пациенты с разделенными полушариями	499
С. Вымышленные мотивы и вымышленные агенты	506
VII. Итоги третьей главы	518
Заключение	521
Библиография	544
Тексты на русском языке	544
Тексты на английской языке	551

Моей дочке Анастасии Волковой

Предисловие

«Папа, у меня появилась идея. Я хочу сделать в школе доклад по твоей философии. Ты можешь мне рассказать про сознание? Что это такое и почему тебя это интересует?» Эта неожиданная просьба застает меня врасплох. Ей всего десять лет, и мне трудно объяснить ей суть философских проблем, которыми я занимаюсь. Мне потребовалось десять лет только для того, чтобы познакомиться с темой, увидеть и сформулировать для себя основные вопросы, разобраться в имеющихся стратегиях ответов на них, ключевых проблемах и аргументах, а теперь мне нужно буквально в нескольких предложениях, да еще и без использования специфической терминологии, самыми простыми словами, изложить все содержание моей работы. И тем не менее я пытаюсь. Я пытаюсь рассказать ей о том, чем я занимаюсь, почему это важно для меня и почему это важно вообще. И эта книга, хотя она, конечно, совсем не похожа на школьный доклад, – это благодарность за ее живой интерес. Я начал писать эту книгу очень давно и для себя, а заканчиваю теперь – для Насти, надеясь, что когда-нибудь она ее прочитает.

Кроме дочки в моей жизни есть и другие важные люди, которые воодушевляли меня во время создания этой книги, без которых она никогда бы не появилась. Это – прежде всего мой наставник, друг и главный собеседник в философии,

Вадим Васильев. Именно он привил мне вкус к философским аргументам, мудро направлял мои исследования, всегда очень деликатно корректировал мои недочеты и на собственном примере, в своих статьях и книгах демонстрировал образец честной и глубокой философской работы. Вместе с Васильевым мы почти десять лет назад создали Московский центр исследования сознания (МЦИС), который сегодня объединил вокруг себя многих специалистов по аналитической философии сознания. Начинался же он с того, что к нам присоединились Роберт Хауэлл, Антон Кузнецов и Артем Беседин.

С их появлением в Центре мы стали настоящим философским сообществом – пусть небольшим, но очень активным. Работать стало намного интересней – появилась возможность дискутировать и получать быстрый живой, эмоциональный отклик. Коллеги – Васильев, Кузнецов, Беседин – слушали мои выступления, читали черновики работ и критиковали предлагаемые мной решения. Благодаря им я уточнил определения, избавился от многих затруднений, лучше определился с собственной позицией и отточил аргументы. Этому способствовали также диалоги с нашими коллегами из Санкт-Петербурга: Данилом Разеевым, Сергеем Левиным, Марией Секацкой, Дарьей Чирвой. Мы познакомились на одной из совместных конференций в Москве, а потом они пригласили нас к себе в Петербург. С этого момента завязалась дружба, и наше профессиональное сообщество расши-

рило свои границы. Впрочем, в связи с работой Центра у меня появилось много собеседников и за рубежом.

Особо продуктивными были международные Летние школы по проблемам свободы воли и моральной ответственности, организованные Центром. Во время этих двухнедельных выездных семинаров мы собирали группы молодых ученых из разных стран, специализирующихся на этой проблематике, и под руководством именитых профессоров и авторитетных специалистов обсуждали основные проблемы и аргументы, полемизировали и делились идеями. На этих Летних школах и различных конференциях я познакомился со многими западными аналитическими философами. Это прежде всего Дерк Перебум, Джон Фишер, Ричард Суинбёрн, Пол и Патрисия Чёрчленд, Джесси Принц и Дэвид Чалмерс. Одно дело, когда читаешь философские работы автора, о котором знаешь только из энциклопедических статей. Совсем другое дело – когда эти авторы становятся твоими собеседниками. Общение с этими философами значительно повлияло на стиль моего мышления. Особую роль в этом сыграл Дэниел Деннет. С ним меня связывает долгая дружба. Для меня Деннет стал образцом современного философа: ходячая энциклопедия (причем по многим отраслям знаний от музыки до биологии), необыкновенно острый математический ум и прекрасный оратор и публицист. Влияние его концепций хорошо прослеживается в тексте всей книги.

В ходе работы над текстом мне удалось защитить доктор-

скую диссертацию на философском факультете МГУ по проблеме свободы воли. На защите возникла острая дискуссия. Не всем членам совета понравился аналитический подход к проблеме свободы воли. Впрочем, как ни странно, во многом успеху способствовали выступления оппонентов Данила Разеева, Светланы Нагумановой и Дмитрия Иванова. Несмотря на то что они не были согласны со всеми выводами работы, они доказали, что аналитическая философия в целом и ее подход к современным метафизическим проблемам заслуживает самого большого внимания. Именно современная аналитическая философия дает лучший образец философского критического мышления и продолжает традиции классического философствования. Я очень благодарен им за оппонирование и критику. Некоторые их замечания я учел при работе над этой монографией.

Кроме того, я благодарен философскому факультету МГУ. Здесь я получал образование, здесь я впервые познакомился с современными философскими вопросами, здесь работают мои учителя и коллеги Дмитрий Бугай, Егор Фалев, Галина Стрельцова, Владислав Коровин и Владимир Миронов. И здесь я воплотил мечту своей юности – стал профессиональным философом. Спасибо за помощь, критику и замечательные дискуссии. И наконец, огромное спасибо моему редактору Андрею Мерцалову, который работал над этим текстом с огромной кропотливостью и усердием иногда в выходные дни и даже по ночам. Благодаря ему мы испра-

вили множество недочетов, и текст читать стало легче. Он был моим верным соратником в этой трудной, некогда казавшейся бесконечной работе. Теперь она закончена.

Введение

Если вол забодает мужчину или женщину до смерти, то вол побить камнями и мяса его не есть; а хозяин вола не виноват...

Библия. Исход. 21: 28

Бартоломе Шасене, французский юрист, практиковавший в XVI в., сделал себе имя в судах, выступая представителем довольно необычных клиентов. Он защищал интересы животных. Наиболее известна его защита крыс, обвинявшихся церковным судом в уничтожении урожая ячменя в епархии Отена. На основании поступившей жалобы было открыто дело, обвиняемые вызваны в суд к определенной дате, а Шасене, талантливый и знающий юрист, назначен их защитником. Причастность крыс была очевидна ввиду множества свидетельских показаний, да и репутация грызунов не была безупречной. Поэтому Шасене пришлось прибегнуть к разным юридическим уловкам, чтобы отстаивать права своих клиентов.

В первую очередь он обратил внимание судьи на то, что подзащитные расселены на значительной территории страны и не были в равной мере проинформированы о вызове в суд. Эту техническую тонкость приняли во внимание, и были сделаны повторные вызовы крыс в суд, на этот раз с кафедр всех церковных приходов епархии. По истечении отпу-

ценного времени Шасене, однако, обратился с новой жалобой. Он заметил, что явка в суд для крыс была сопряжена с риском для жизни, так как на каждом шагу их подстерегали злейшие враги – кошки. Адвокат утверждал, что любой имеет право воздержаться от появления в суде, если это угрожает его безопасности. Это обстоятельство также приняли во внимание, благодаря чему адвокату удалось избежать дефолта из-за неявки клиентов. Впрочем, окончательный вердикт этого трибунала остается неизвестным – он не отражен в доступных источниках. Известно, однако, что именно этот процесс принес Шасене славу и способствовал его карьере, что в итоге позволило ему занять пост первого председателя высшего суда Прованса. Известно и то, что проблема защиты животных в судах интересовала Шасене и дальше. Он даже написал на эту тему объемную работу. Основным ее выводом служило то, что животные могут представлять перед судом, однако в процессе этих судов должны выполняться все юридические нормы и правила. До тех пор пока эти процедуры не соблюдены, к обвиняемым не могут применяться никакие наказания. Мнение Шасене о наличии вины и подсудности животных разделяли и некоторые другие средневековые юристы.

В исторической хронике, созданной швейцарским летописцем Шиллингом, сохранился отчет о суде, состоявшемся в 1479 г. под председательством епископа Лозанны. Дело было открыто в связи с жалобой местного приходского свя-

щенника на неких насекомых, называемых *inger* (по-видимому, зерновки). Согласно свидетельству священника, насекомые уничтожили значительную часть урожая. Епископ не остался равнодушным к описанию преступлений, совершенных вредителями, и издал указ, в котором наказывал им в шестидневный срок покинуть поля, сады и деревья. В случае неповиновения насекомые должны были предстать перед судом города Вифлисабурга (совр. Аванш).

Из хроники следует, что насекомые не испугались указа. На судебных заседаниях интересы насекомых представлял адвокат Жан Перродэ из Фрибурга. Он не был таким же изобретательным, как Шасене, поэтому суд завершился достаточно быстро и не в пользу обвиняемых. Судья счел преступления доказанными, а вину очевидной. В качестве наказания насекомые были преданы анафеме, и им было запрещено впредь появляться на полях. Впрочем, анафема не была самым жестоким наказанием, применявшимся к животным. По итогам множества средневековых судебных разбирательств и судов Нового времени животных нередко пытали, отправляли в тюрьмы, вешали, душили, сжигали и закапывали заживо.

Так, за убийство ребенка в 1456 г. в городке Оппенхайм и в 1557 г. в коммуне Сен-Кантен на смертную казнь были осуждены свиньи – их похоронили заживо. В 1314 г. бык, принадлежавший фермеру в селении Муази, на дороге напал на мужчину. Мужчина погиб от ранений, а животное было

поймано и посажено в тюрьму, после чего состоялся суд, на котором быка приговорили к повешению. В 1474 г. в Базеле на смерть за необычное преступление был осужден петух – он якобы снес яйцо. Это преступление было приравнено к колдовству, и птицу сожгли заживо. Американский историк Э. Эванс приводит больше сотни таких случаев за XVI–XVII вв. Основанием для таких действий, вероятно, было убеждение в универсальности закона – так называемого *lex talionis*, закона равного возмездия. Удивительным образом он распространялся даже на неодушевленные предметы.

В Древней Греции, например, существовала специальная категория разбирательств – суды над вещами. Так, согласно Драконовым законам в Древней Греции оружие и любые предметы, задействованные в убийстве, должны были быть публично осуждены и выброшены за пределы Афин. Из Античности известен случай осуждения статуи атлета Никона с Тасоса. Подвергшись нападению завистников спортсмена, статуя была свергнута с пьедестала, но, падая, обрушилась на одного из нападавших. В результате статуя была представлена перед трибуналом и осуждена на утопление в море.

Любопытна в этом отношении и знаменитая расправа над угличским колоколом. С помощью этого колокола в 1591 г. жители города были оповещены о гибели царевича Дмитрия. Угличане расправились с предполагаемыми убийцами наследника престола, а царь Борис Годунов решил наказать их за самосуд. Наказали не только людей, но и колокол: его

сбросили со Спасской колокольни, вырвали ему язык, отрубили ухо, принародно на площади он был избит плетьюми, а после – отправлен в сибирскую ссылку. В Тобольске воевода, князь Лобанов-Ростовский, велел запереть колокол в приказной избе и сделать на нем надпись «первосыльный неодушевленный с Углича». Об этом свидетельствуют «Сибирские летописи» и «Статейный список сибирских воевод».

Однако наиболее жутким примером работы системы правосудия был «трупный синод», состоявшийся в январе 897 г. в Латеране. Этот трибунал был проведен над эксгумированным трупом папы Формоза, который умер в апреле 896 г. В Латеранской базилике Св. Иоанна новый папа Стефан велел посадить на трон полуразложившийся труп своего предшественника и в присутствии высшего духовенства подверг его посмертному допросу. Формозу вменялось вероломство, переход с одной епископской кафедры на другую вопреки установленному Никейским собором запрету, а также совершение им, мирянином, религиозных таинств. Итогом разбирательства стало признание указов бывшего папы недействительными, и... наказание тела. У трупа были отрублены три пальца, которыми Формоз совершал крестное знамение. После чего его проволокли по улицам Рима и с прикрепленным грузом сбросили в Тибр.

В некотором смысле подобные суды носили символический характер. Их целью была в том числе демонстрация

универсальности закона для всех и всего, и в этом смысле их можно обосновать. Но, судя по всему, одной из мотиваций таких судов было возмездие. И в этом смысле подобные процессы абсурдны. Возмездие, как и вознаграждение, можно применять только к объектам, которые могут нести ответственность. Только такие объекты способны совершать преступления и быть виновными, поступать хорошо и быть достойными благодарности. Этим требованиям иногда соответствуют люди. Но животные и неодушевленные предметы, скорее всего, для этого не подходят. Объяснить причину такого несоответствия можно, выявив условия ответственности.

Понятие ответственности родственно понятию причины. В связи с этим иногда говорят, что «лавиная несет ответственность за гибель альпинистов», или что «ураган ответственен за многочисленные разрушения». Но это – переносное значение слова. Ответственность в прямом смысле, с точки зрения современной юридической практики и большинства современных этических теорий, – это нечто большее. Она предполагает выполнение особых условий для агента и поступка. Этими условиями являются зависимость действий от намерений, наличие свободы воли и какой-то особой связи, сохраняющейся во времени между агентом и совершенным действием.

Во-первых, ответственность предполагает, что действие в качестве своего источника имеет намерение. События, не

связанные с намерениями агентов, как и случайные события, не влекут за собой обвинения или благодарность. Конечно, существуют, например, преступления, совершенные неумышленно. Однако в таком случае преступный поступок является следствием другого, намеренного действия, которое привело к преступлению. Так, если пьяный водитель совершает наезд на прохожего, мы считаем водителя виновным, даже если он совершил наезд непреднамеренно. Здесь вина проистекает из того, что водитель намеренно употреблял алкоголь и после этого сел за руль. Таким образом, намеренное действие все равно является необходимым в цепочке ответственного действия.

Вторым условием ответственности является свобода воли, понимаемая в первую очередь как возможность поступать иначе. Если возможности поступать иначе у агента нет, его действие считается вынужденным. Но вынужденные действия обычно не относятся к категории тех, за которые агент может нести ответственность. Так, за действия, совершенные под дулом пистолета или под угрозой расправы над личностью или его родственниками, человек обычно не несет ответственность. В подобных случаях ответственность возлагается на тех, кто вынудил совершить действие, и снимается с выполнившего чужое распоряжение агента. Только за действия, продиктованные свободным волеизъявлением, человека можно считать ответственным.

Наконец, необходимо, чтобы агент, совершивший дей-

ствие, был особым образом связан с действием, за которое ему полагается воздаяние. В качестве этой связи подходит отношение принадлежности: действие должно было принадлежать тому же агенту, кому полагается воздаяние. Связь с другими агентами не подходит для сохранения ответственности. Очевидно, например, что если преступление совершил один близнец, то его брат не несет ответственность. Не может нести ответственность сын за поступки своего отца. Конечно, сын может испытывать угрызения совести или осуждать поступки отца, но это не то же самое, что ответственность. Оснований наказывать или поощрять сына за поступки отца нет.

Таковы три важнейших условия ответственности. Возможно, существуют и другие условия, но я сфокусирую внимание именно на этих трех. Эти три условия не выполняются в случаях, о которых шла речь выше, в ситуациях с неодушевленными предметами и нерациональными существами. Однако все эти условия должны быть выполнимы в принципе и фактически выполняться хотя бы иногда в жизни нормальных личностей. В противном случае мы бы лишились оснований приписывания моральной ответственности, а вместе с ней – и того фундамента, на котором строится большая часть межличностных отношений.

События у людей вызывают различные реакции, положительные или отрицательные. Люди радуются или огорчаются тому, что пошел дождь. Они сожалеют, что дерево упа-

ло на машину, или довольны, что в этом году поспело много яблок. Но есть специфические реакции, которые применяются только к определенному роду событий – ответственным человеческим действиям. Этими специфическими реакциями является негодование, осуждение, благодарность и т. п. На дождь или дерево нерационально злиться и негодовать. Неуместно благодарить или восторгаться поведением яблонь и светофоров. Негодовать, злиться, испытывать благодарность можно только в отношении ответственных действий агентов.

Представьте, что эти реакции исчезли бы совсем: каким был бы мир без них? Отношения между людьми стали бы полностью «объективными». К примеру, с человеком, совершившим дурной поступок, поступили бы так, словно он – упавшее на машину дерево или плохо работающий пылесос. Его бы починили, исправили, привели в пригодное состояние. Но по отношению к нему не было бы гнева и осуждения. Да и сам человек не мог бы испытывать раскаяния. Был бы такой мир лучшим? Мнения об упразднении отрицательных ответных реакций могут расходиться. Кто-то может считать, что, возможно, было бы даже очень неплохо, если бы в отношениях между людьми исчезли гнев, жажда мести и даже чувство вины. Но упразднение положительных ответных реакций – думаю, с этим согласятся все – сделало бы отношения бессодержательными, утилитарными. Можно ли помыслить отношения без чувств благодарности, симпатии, люб-

ви, уважения? Нельзя! Эти чувства принципиально важны и наполняют жизнь смыслом. Но все эти реакции продиктованы представлением об ответственности. И можно попытаться показать, что условия ответственности невыполнимы в принципе.

Существует точка зрения, что признание нормальных здравомыслящих личностей ответственными так же необоснованно, как и приписывание вины животным и неодушевленным предметам. То есть предполагается, что действия и поступки людей принципиально не отличаются от любых других событий в природе. Это мнение поддерживается некоторыми современными научными и философскими теориями. Может показаться, что среди научных теорий опасность для ответственности должны представлять социально-психологические и нейрофизиологические теории. Так, против ответственности могут быть использованы теории, в которых доказывается определяющая роль ситуации или авторитета в поведении людей. Доводом может служить, в частности, знаменитый Стэнфордский тюремный эксперимент, проведенный в 1971 г. психологом Филипом Зимбардо¹.

¹ Эксперимент является одним из классических в современной социальной психологии. Однако он часто критикуется как в этическом плане, так и в плане соблюдения научной методологии. Некоторые критики считают, что Зимбардо фактически доказал не то, что поведение личности определяется ситуацией, а то, что конформность с авторитетом определяют поведение. Однако для целей нашего исследования это не имеет существенного значения. Эксперимент все

Эксперимент начался с объявления в газете, в котором молодым мужчинам предлагалось за небольшую плату две недели прожить в тюремных условиях в роли «охранника» или «заключенного». «Тюрьма» была создана на территории Стэнфордского университета, в подвале здания психологического факультета. Помещение состояло из коридора и нескольких небольших кабинетов. Часть из них была превращена в камеры. Из комнат вынесли мебель, поставили простые деревянные кровати, двери заменили решетками. Чулан в коридоре стал одиночным изолятором. В конце коридора несколько дополнительных комнат отвели для охранников и наблюдателей. Во всех помещениях было смонтировано аудио- и видеонаблюдение, так чтобы все события фиксировались.

Роли «заключенных» и «охранников» были распределены случайным образом. Те, кому выпала роль «охранников», должны были работать в три смены по 8 часов. После работы они возвращались домой. «Охранникам» объяснили, что целью эксперимента является симуляция тюрьмы в рамках практических и этических ограничений. Главной их задачей было поддержание порядка и дисциплины. Для этого, в частности, они должны были регулярно пересчитывать «заключенных». Никто не должен был пропасть. «Заключенным» специальных инструкций не давали. Предполагалось, что все время они проведут в камерах. Может быть, именно из-за от-

сутствия обязанностей большее число участников изначально изъявляли желание быть именно «заключенными». Впрочем, энтузиазм тех, кому выпала эта роль, скоро убавился. Организаторы постарались, чтобы условия эксперимента не сильно отличались от реальности.

Утром первого дня экипаж полицейских произвел задержание «подозреваемых». Это оказалось довольно неожиданным. Будущим «заключенным» были предъявлены обвинения, на них надели наручники и с завязанными глазами на полицейской машине привезли в помещение «тюрьмы». Там «охранники» по протоколу провели прием поступивших: их обыскали, раздели догола, обрызгали из баллончика с дезодорантом, якобы для дезинфекции, одели в новую тюремную форму. Этой формой были робы с номерами, под которыми запрещалось что-либо носить. На щиколотку «заключенным» надели цепочку с замком, а волосы закрыли нейлоновой шапочкой. Выполнение процедур предназначалось для того, чтобы придать еще большего реализма тюремной обстановке, а форма и присвоение номеров способствовали «деперсонализации» «заключенных».

«Деперсонализации» были подвергнуты и «охранники». Им были выданы одинаковые рубашки цвета хаки, деревянные полицейские палки, свистки и черные очки. Эти атрибуты носили символический характер. Палки применять запрещалось, однако все это должно было способствовать вживанию волонтеров в свои роли, а также формированию но-

вой идентичности. Как показал эксперимент, это произошло даже быстрее и радикальней, чем могли предположить организаторы. Пожалуй, наиболее явно это стало проявляться во время «пересчетов».

Эта рутинная процедура формально предназначалась для проверки наличия всех «заключенных» на месте. Первые пересчеты также предназначались для ознакомления заключенных со своими номерами. Но «охранники», ощутив свою власть над «заключенными», превратили процедуру в психологическое истязание. С шутками и издевками они заставляли «заключенных» десятки раз повторять номера, в разных порядках, с разными интонациями, пропевая их, делая различные жесты. «Заключенные» недоумевали. Однако любые попытки сопротивляться этим издевательствам стали вызывать агрессию. «Охранники», освоившись со своей ролью, осознавшие собственную безнаказанность, заставляли десятками раз повторять одно и то же. Тех, кто не выполнял команды или делал это неохотно, раздевали догола, заставляли приседать или отжиматься. Через некоторое время и это стало казаться недостаточным: «охранники» стали ставить на отжимающихся ноги или заставляли других «заключенных» садиться сверху. Пересчеты, которые должны были длиться десять минут, стали растягиваться на часы.

Издевательское поведение «охранников» вызвало протест среди «заключенных». Им было трудно смириться с подобным произволом, ведь в их представлении «охранники» не

отличались от «заключенных»: и те и другие были просто участниками психологического эксперимента, и роли были определены случайным образом. На второй день мятежники забаррикадировали собственные камеры и отказались выходить. Но сопротивление вызвало еще большую агрессию со стороны «охранников». Мятеж был подавлен с крайней жестокостью. «Охранники» вызвали дополнительную смену в подмогу и стали распылять огнетушители через решетки в камеры. Задыхающиеся «заключенные» вынуждены были сдаться, но им не удалось избежать наказания. Наиболее непокорных заключили в изолятор, у остальных забрали кровати и постельное белье. Но даже на бетонном полу им не удавалось выспаться.

«Охранники» стали проводить пересчеты и ночью с новыми и новыми издевательствами. Нервные срывы и беспомощность «заключенных», казалось, вызывали еще больше энтузиазма. «Охранники» стали раньше приходиться на работу и придумывали все больше и больше бесполезных и абсурдных действий, которые они вынуждали совершать «заключенных» (например, заправлять и разглаживать постель, перетаскивать с места на место пустые коробки и т. п.). Более того, они запретили бунтарям пользование туалетом – это стало привилегией наиболее послушных «заключенных». Остальным в камеры поставили ведра, так что вскоре запах и гигиеническое состояние «тюрьмы» стало совершенно безобразным. Когда издевательства стали носить порно-

графический характер, руководитель эксперимента вынужден был вмешаться. За пять дней молодые, психически здоровые ребята из благополучных семей, никогда не проявлявшие склонность к насилию, превратились в садистов. Как это могло случиться? Откуда возникла такая изощренность и изобретательность в способах унижения?

По мнению Ф. Зимбардо, основной причиной этого была Ситуация. Сама сущность «тюремного порядка» и распределение ролей оказались достаточным основанием для возникновения аберрантного, антисоциального поведения «охранников». Введенные символические элементы и процедуры, из которых состоял «порядок», спровоцировали агрессию. В этих условиях индивидуальный моральный характер оказался подавленным. В очень короткий срок внешние обстоятельства трансформировали личности даже тех, кто не имел никаких первоначальных склонностей к садизму. Таким образом, эксперимент показал, что тюремный порядок ведет к моральной деградации стражей. Однако, как заметили наблюдатели, моральная деградация распространяется и на заключенных. С течением времени даже они начинают копировать поведение стражников – отношение заключенных друг к другу также зависит от общей динамики. То есть все участники, независимо от их роли, подпадают под контроль ситуации и это ведет к моральной деградации участников. Из эксперимента следует, вероятно, что тюремный порядок не соответствует поставленным обществом целям.

Впрочем, из Стэнфордского тюремного эксперимента, как я уже говорил, можно сделать и более общий вывод. Возможно, из него следует, что внешние обстоятельства определяют действия индивидов и что разница между преступниками и жертвами – только в ситуации, в которой они оказываются. Одним выпадает роль преступников, другим – стражей, и в этом кроется объяснение различий. Подобный вывод можно попробовать обосновывать и результатами другого знаменитого эксперимента. Его провел один из самых влиятельных психологов XX в., Стэнли Милгрэм из Йельского университета. Милгрэм показал, что под воздействием авторитета люди способны совершать аморальные поступки, которые они никогда не совершили бы самостоятельно.

В Эксперименте о подчинении 1963 г. участвовали сорок мужчин в возрасте от двадцати до пятидесяти лет из разных социальных групп. Среди них были инженеры, школьные учителя, продавцы, рабочие и даже ученые. Участникам объяснили, что цель эксперимента в том, чтобы исследовать влияние боли и наказаний на память. Эксперимент начинался с распределения ролей: участники тянули жребий и, как они думали, могли получить роль либо учителя, либо ученика. На самом деле все участники получали роль учителя. Роль ученика всегда выпадала приглашенному актеру.

По формальным условиям эксперимента «ученик» должен был запоминать пары слов из длинного списка. После чего его на глазах «учителя» привязывали к креслу с элект-

тродами «для того, чтобы ограничить нежелательные движения». «Учитель» переходил в другое помещение и садился за стол с контрольной панелью, которая должна была управлять электрическим генератором. Панель состояла из 30 переключателей. На каждом переключателе были подписаны значения вольт от 15 до 450. Переключатели были объединены в несколько групп: «легкий шок», «умеренный шок», «сильный шок», «очень сильный шок», «чрезвычайно сильный шок», «опасность: максимальный шок». Согласно условиям эксперимента, «учителя» должны были проверять знание списка у «ученика» и в случае неправильного ответа наказывать «ученика» электрическим разрядом. За каждый последующий неправильный ответ уровень шока должен был повышаться.

Перед испытанием эксперты сделали прогноз, что не более 3 % участников будут «наказывать» «учеников» электрическим шоком с высокими значениями. Их ожидало ужасающее открытие: почти все участники в конечном итоге использовали шок в 450 вольт. То есть почти все здоровые здравомыслящие люди готовы были подвергнуть «ученика» сильнейшим страданиям в условиях лабораторного эксперимента. Может быть, они не понимали, что должно было происходить, или не верили в реалистичность эксперимента? Может быть, они подверглись принуждению со стороны организаторов?

«Ученик» в действительности не испытывал никакого шо-

ка – оборудование лишь симулировало электрические разряды. Но эксперимент был поставлен так, что участники никаким образом не могли догадаться об этом. Для убедительности «учителям» вначале самим предлагалось протестировать аппарат, и каждый из них получал довольно сильный разряд в 45 вольт. Они также слышали реакцию «ученика»: после шока в 300 вольт «жертва» начинала стучать ногами в стену, требуя прекратить эксперимент. То же повторялось и после разряда в 315 вольт. Перед каждым использованием переключателя участник должен был вслух объявлять значение электрического разряда. То есть участники были уверены, что причиняют боль. Это подтверждалось и их реакцией. Большая часть испытуемых были крайне эмоционально напряжены. Они краснели, кусали губы, тяжело дышали, дрожали, царапали себя ногтями и тем не менее продолжали эксперимент. При этом никакого физического принуждения они не испытывали. У них вообще не было обязательств продолжать «пытку».

В комнате с испытуемым находился ведущий эксперимента. Он объяснял цели исследования и его правила. Когда испытуемый «учитель» останавливался (например, после стука в стену и протестов «ученика»), ведущий говорил, чтобы он продолжал. Когда от «ученика» вообще переставали поступать сигналы, ведущий указывал на то, что отсутствие ответа должно приравниваться к неправильному ответу и что за это должно следовать «наказание». Среди стандартных заго-

товленных фраз ведущего на возражения со стороны «учителей» были: «пожалуйста, продолжайте», «эксперимент требует, чтобы вы продолжили», «абсолютно важно, чтобы вы продолжали», «у вас нет иного выбора, кроме как продолжать». Но никаких других способов воздействия на испытуемого у ведущего не было. Участник в любой момент мог просто прекратить пытку. Это теоретически мог сделать каждый испытуемый, при этом большинство из них продолжали до самого конца, вопреки убеждению, что причинение боли другому является злом. Почему?

По мнению Милгрэма, в данном случае определяющими факторами были авторитет и склонность подчиняться. Участники эксперимента понимали, что он проводится таким авторитетным учреждением, как Йельский университет, они предполагали, что его сотрудники компетентны и ответственны за условия, и поэтому, вопреки внутренним моральным соображениям, следовали всем указаниям ведущего. Удивительным было лишь то, что этот авторитет продолжал играть решающую роль даже тогда, когда действия начинали носить абсолютно аморальный характер. Они нарушали прививаемую всем с детства максиму о непричинении боли и страдания другому. И при этом участники действовали в отсутствие внешнего принуждения. Несмотря на то что неподчинение ведущему эксперимента не грозило никакими последствиями испытуемым, они почти все избрали путь конформизма. Они выбрали подчинение даже несмотря на

то, что испытывали сильнейшие угрызения совести и эмоциональный стресс. Таким образом, описанный выше эксперимент демонстрирует преобладание силы конформизма и привычки к подчинению над моральными ценностями личности.

В этом отношении выводы Милгрэма созвучны результатам, полученным Зимбардо: внешние обстоятельства играют большую роль, чем мы ожидаем. Они во многом определяют поведение личностей. Таким образом, описанные выше эксперименты могут заставлять усомниться в выполнимости условий ответственности. Впрочем, наука выявляет и другие, *внутренние* обстоятельства, которые играют важнейшую роль в определении поведения – это неосознаваемые процессы, протекающие в мозге. Нейрофизиология раскрывает механизмы, определяющие поведение человека. Одной из лучших иллюстраций в этой области является эксперимент Бенджамина Либета 1983 г.

Ученый предложил испытуемым двигать пальцем и/или запястьем правой руки тогда, когда им этого хотелось. Движение не должно было зависеть от каких-либо внешних факторов, но совершаться в произвольный момент. Участников предупредили, что они не должны планировать движение заранее. При этом нужно было точно фиксировать момент принятия решения о движении руки. Для этого использовались специальные точные часы: световая точка на циферблате осциллоскопа пробегала полный круг каждые примерно

2,5 секунды, поэтому с помощью положения точки на циферблате можно было определить время с точностью до доли секунды. С помощью этих часов испытуемые фиксировали момент возникновения желания пошевелить рукой. Но параллельно с этим Либет записывал еще один параметр – потенциал готовности.

К голове над двигательной областью коры головного мозга испытуемым прикрепляли электроды. С помощью этих электродов безопасным для участников способом ученые снимали показания об изменении электрической активности в мозге и возникновении потенциала готовности, который всегда предшествует двигательной активности, сокращению мышц. Время возникновения потенциала готовности также записывалось для каждого «произвольного» действия, а потом сравнивалось с временем фактического действия и моментом возникновения желания, фиксируемым самим участником эксперимента. Как соотносились эти показания по времени?

Логичным было бы предположить, что первым событием должно было бы быть возникновение желания, фиксируемое самим испытуемым с помощью часов, ведь по условиям эксперимента решение о движении руки должно было быть произвольным и осознанным. Следом стоило бы ожидать возникновения потенциала готовности, так как он необходим для инициации действия. И наконец, само движение должно было замыкать каузальную цепочку. К удивлению уче-

ных, последовательность оказалась другой. У испытуемых первым возникал потенциал готовности, и только затем следовала фиксация желания. То есть к тому моменту, как у испытуемого возникало желание совершить движение, его мозг уже принял это решение. Это открытие может быть еще одним доводом в пользу утверждения об отсутствии у агента свободы, а значит, и ответственности. Из некоторых интерпретаций этого эксперимента следует, что свобода воли – это не более чем удобная иллюзия.

Из описанных выше экспериментов может следовать, что поведение личности определяется факторами, лежащими вне ее контроля: социальными силами и нейрофизиологическими процессами. Социальная динамика направляет личность, и поведение оказывается обусловлено не контролируемыми агентом факторами. Сознание лишь регистрирует решение, устанавливаемое неосознаваемыми процессами в мозге, а агент испытывает иллюзию выбора. Но если это так, то три описанных выше условия ответственности не выполняются даже для здоровых нормальных личностей.

Эксперименты Зимбардо, Милгрэма и Либега, как кажется, демонстрируют, во-первых, что намерения агентов не являются источником их действий – этим источником служат нейрофизиологические бессознательные процессы или социальный контекст. Во-вторых, они, вероятно, показывают, что социальный контекст определяет поведение личности таким образом, что у нее стираются индивидуальные особен-

ности, то есть у личности не остается альтернатив. И в-третьих, они, возможно, свидетельствуют, что подлинным субъектом действия является мозг человека, его неосознаваемая активность, а не сознательная личность, выступающая объектом воздаяния и наказания. Таким образом, ни одно из указанных условий не выполняется, а если это так, то никто и никогда не должен нести ответственность, и наказание и благодарность в отношении людей столь же абсурдны, как наказание и благодарность в отношении животных или даже неодушевленных объектов. Следует ли принять такой вывод на основании эмпирических данных? Приводит ли наука к такому заключению?

Голландский философ и натуралист Б. Спиноза считал, что люди верят в свободу воли потому, что осознают собственные желания, но не осознают реальные причины этих желаний. По этой логике, сейчас, после прояснения некоторых причин поведения, у свободы воли и связанной с ней ответственности не остается оснований. Такую позицию разделяют некоторые выдающиеся современные западные ученые, среди которых Д. Вегнер, С. Харрис, М. Газзанига, Б. Либет, Ф. Крик и др. Есть и российские ученые, занимающие подобную позицию, например В. В. Никулин и В. А. Ключарев. Я, однако, не согласен с такого рода воззрениями. Я считаю, что любые эмпирические открытия сами по себе не могут доказать отсутствие или наличие ответственности и свободы. Открытия новых, ранее неизвестных и неосознавае-

мых причин действий не представляют действительной угрозы для этих важных убеждений. Эти причины могут представлять препятствие для ответственности только в составе философских теорий. Почему научные теории сами по себе не показывают, что свободы не существует? Почему ответственность и необходимая для нее свобода воли – это философские проблемы?

Предположим, эксперименты показывают, что источником действий личностей действительно являются неосознаваемые нейрофизиологические процессы. Означает ли это, что в таком случае намерения не могут быть источником действий? Ответ на этот вопрос лежит явно за пределами нейрофизиологии, ведь для этого нужно соотнести понятия «намерение» и «неосознаваемый нейрофизиологический процесс». «Намерение» входит в язык бытовой психологии и не имеет строгого определения. Чтобы произвести соотношение с терминами нейрофизиологии, нужно сначала прояснить его с помощью концептуального анализа, осуществить перевод, расшифровку. Это – задача для современных философов, занимающихся проблемой сознания. Выходит, чтобы выяснить, возможна ли свобода, нужно решить вопрос о соотношении ментального и физического – одну из ключевых философских загадок, разрешить многовековой философский спор. Но даже после этого решения философская работа не заканчивается – нужно разработать теорию причинности.

Предположим, намерение и его нейронный коррелят различаются между собой. Это по-прежнему не гарантирует, что оба явления одновременно не могут быть источниками действия. Источник действия – это причина действия, но именно философия изучает свойства причин. Может ли у явлений быть одновременно несколько причин, могут ли эти причины систематически коррелировать – все это вопросы теории причинности. Они должны быть определенным образом решены, чтобы понять, действительно ли неосознаваемые нейронные процессы являются препятствием для свободы воли. Впрочем, и само представление о свободе требует философского исследования. Что под ним следует подразумевать?

Согласно распространенному мнению, свобода предполагает наличие альтернативных возможностей. Это правдоподобное утверждение, но что именно здесь предполагается? Мы не можем непосредственно наблюдать альтернативные возможности, живя в мире актуальном. Тогда как определить их наличие или отсутствие? Здесь опять нам приходится прибегать к философии. Эмпирическая наука изучает природу преимущественно в том виде, в котором она существует, а альтернативные возможности – это гипотетические ситуации, «что, если бы...» ситуации. Мир гипотетических возможностей – это мир философа. Существует множество философских теорий, в которых разъясняется, что значит быть возможным, существовать как реальная альтерна-

тива. Философы различают логические и естественные возможности. Некоторые утверждают, что альтернативных возможностей вообще не существует, а некоторые – что они так же актуальны, как и фактический мир. Это значит, что, даже зная, что все поведение человека диктуется ситуацией, нельзя утверждать отсутствие у него альтернатив *tout court*.

И наконец, третий вопрос – об особой связи действующего агента и агента, несущего ответственность, – вопрос тоже философский. Для ответа на него нужно в первую очередь прояснить, какого рода связь необходима, а потом определить, может ли она быть реализована. Некоторые считают, что действующий и несущий ответственность агент должны быть тождественны. Отношение тождества в различные моменты времени является философским понятием, так как относится к наиболее общему типу отношений. И именно это отношение необходимо прояснить, прежде чем делать выводы о возможности агента быть ответственным. Более того, для отождествления действующего и несущего ответственность агента необходимо понять, что вообще следует считать агентом, какого рода эта сущность и из чего она может состоять. Таким образом, даже имея эмпирические доказательства, что мозг или какая-то его часть является действующим началом, а ответственным – целостная сознательная личность, без философского анализа нельзя утверждать, что действующее начало и личность не совпадают. Все эти соображения приводят нас к выводу о необходимости уча-

ствия философии в решении вопроса об ответственности и свободе. К сожалению, порой философским анализом пренебрегают. И даже выдающиеся ученые, совершившие значимые открытия в науке, не освоив философской методологии, делают ошибки, свойственные неофитам. Методология, которую я считаю наиболее эффективной, – это методология аналитической философии.

Истоки аналитической философии восходят к диалектике Платона в древности и размышлениям Декарта и Гоббса в Новое время. Наиболее влиятельными фигурами, сформировавшими аналитическую философию XX в., считаются Б. Рассел и Л. Витгенштейн. Оба этих философа стремились добиться максимальной ясности и математической точности в своих размышлениях. Эти принципы нашли отражение в ключевом методе аналитической философии – концептуальном анализе. Это исследование сложных понятий с помощью разложения их на более простые. Философский концептуальный анализ близок научному анализу. Так, развитие физики выражается, в частности, в возможности декомпозиции атома на протоны, нейтроны и электроны, развитие химии – в разложении сложных химических соединений на простые элементы. Другим методом аналитиков является эксперимент. Как и ученые, современные аналитические философы ставят эксперименты, но эти эксперименты – не физические, а мысленные. Мысленные эксперименты позволяют выявить базовые интуиции и показать, как они работа-

ют в той или иной гипотетической ситуации.

Еще одной важной чертой аналитической философии является ее связь с эмпирическими данными. Современная аналитическая философия не существует в изоляции, а является фундаментом для междисциплинарных обобщений. Это взгляд на научные открытия с высоты птичьего полета. Необходимость такой перспективы я показал выше. В этой книге я пытаюсь применить комплекс методов аналитической философии к проблемам ответственности и свободы воли.

Цель размышлений, изложенных в этой книге, – решение загадок о свободе воли и ответственности. В этом тексте я попытаюсь проанализировать наиболее значимые философские позиции и ответить на вопрос, как возможны свобода и ответственность. Выше было показано, что одни только эмпирические открытия, научные эксперименты, раскрывающие причины поведения человека, не способны убедить в отсутствии этой свободы и связанной с ней ответственности. Но в совокупности с некоторыми философскими теориями подобные научные открытия могут представлять подлинную опасность для убеждений в наличии у нас ответственности и свободы воли. Такими теориями являются физикализм, детерминизм и событийный каузализм. Каждой из этих теорий и диалогу с ее сторонниками в современной философии посвящена отдельная глава книги. В каждой из глав представлены альтернативные позиции по отдельной философ-

ской проблеме и намечено ее решение. В первой главе обосновывается возможность сочетания физикализма и ментальной каузальности, во второй – детерминизма и возможности альтернативных действий, в третьей – событийного каузализма и возможности приписывания действий личности. Итогом книги является доказательство возможности свободы и моральной ответственности, то есть вывод о том, что события нашей жизни – не просто ирония судьбы.

Неискушенный читатель может встретить в книге много незнакомых терминов. Но пусть это не вызывает удивления. Философия, как и любая другая наука, имеет свой словарь. Он необходим для точности выражения мысли. Но этот лексикон достаточно легко освоить. Сложнее с понятиями, встречающимися как в философии, так и в обыденном языке. У философов эти понятия часто имеют специальное значение (например, понятия «сознание», «причина», «агент»). В книге я не всегда разъяснял их, так как писал этот текст в основном для профессиональных философов, работающих в аналитической традиции. Надеюсь, впрочем, что большая часть терминологии будет понятна из контекста и чтение доставит вам удовольствие, даже если это – первая философская книга на тему свободы воли, которую вам доводилось держать в руках.

Глава 1. Проблема ментальной каузальности и возможность действий

I. Аргументы в пользу ментальной каузальности

1. Связь проблем ментальной каузальности и свободы воли

Ответственность предполагает существование свободы воли и возможность действий. Действия не существовали бы без ментальной каузальности – способности ментальных состояний, убеждений, желаний, эмоций быть причинами поведения. Ведь действия отличаются от обычных событий в силу этих особых причин. Но могут ли ментальные состояния быть причинами поведения? Обосновано ли различие между действиями и обычными событиями? Обладаем ли мы свободой и есть ли основание у ответственности за поступки? Как ни странно, это не пустые риторические вопросы. В начале этой главы я расскажу о главном основа-

нии для сомнений в ментальной каузальности – физикализме. Физикализм – это убеждение в том, что все события в качестве своих причин имеют только другие физические события. Это фундаментальный принцип современной науки. При этом физикализм создает серьезную угрозу для возможности существования ментальной каузальности. Кажется, утверждение о том, что только физические события могут быть причинами любых других событий, ставит под сомнение возможность ментальной каузальности. Эта интуиция выражается в одном из важнейших аргументов современной аналитической философии – Аргументе исключения.

Из Аргумента исключения следует невозможность ментальной каузальности, отрицание свободного и ответственного действия. Такое следствие противоречит интуициям и показывает, что наше отношение к поступкам других людей иррационально. Неужели люди в принципе не могут поступать свободно? Неужели нет оснований для наказаний и похвалы? Надеюсь, что это не так. В поисках решения проблемы ментальной каузальности, разрешения противоречия между нашими интуициями и фундаментальным научным принципом, я обращусь к наиболее значимым, по моему мнению, теориям современных аналитических философов: локальному интеракционизму, биологическому натурализму и телеофункционализму. После исследования этих подходов я представлю собственное дополнение к одному из

них – телеофункционалистскому решению проблемы ментальной каузальности. Итогом главы будет телеофункционалистское обоснование возможности ментальной каузальности в условиях физикализма. Таким образом, будет преодолено первое препятствие на пути к возможности свободы и ответственности. Другие я рассмотрю во второй и третьей главах. А теперь сделаем шаг назад и попробуем убедиться в том, насколько действительно ментальная каузальность необходима, ведь само по себе это не очевидно.

2. Корреляция ментальных состояний и действий

Среди всего множества процессов и событий можно выделить те, что являются действиями. Особенность действий в том, что они предполагают наличие агента, действующего активного субъекта. Действия каким-то образом связаны с агентом и его психическими состояниями. Но какова эта связь и чем различаются события в жизни агента и его действия? Американский философ Д. Дэвидсон задается подобными вопросами в эссе «Агентность»: «В каких событиях в жизни личности проявляется агентность? Что является ее действиями и поступками в отличие от того, что просто случается в ее истории? Какая отличительная черта характеризует ее действия?» [Davidson 1980, 43]. Сегодня я проснулся рано утром, разбуженный звуком скрипки, на которой кто-то

занимался, умылся, побрился, налил себе чашку кофе, споткнулся о ковер и разлил кофе по неуклюжести. Часть из того, что случилось, – мои действия, а часть – просто события из моей жизни. Так, пробуждение и спотыкание – события, а остальное – действия. Но чем они отличаются?

Кажется, действиями являются те события, которые объясняются внутренними состояниями агента. И кажется, что внутренние состояния агента объясняют действия. Это утверждение разделяют многие аналитические философы, например А. Макинтайр. В монографии «После добродетели: исследование теории морали» он пишет: «Определить событие как действие означает... дать ему такое описание, чтобы это событие разумно и логично проистекало из намерений, мотивов, стремлений и целей агента-человека» [MacIntyre 2007, 209]. От других событий действия отличаются тем, что у них есть описание, логически проистекающее из психических состояний агента.

Вот в чайнике закипела вода. Кто-то поставил его на плиту и включил конфорку. Процесс разогревания чайника – не только ряд физических событий, но и действие агента. Конечно, нагревание и кипячение проходят по законам термодинамики. Но в отличие от других событий и процессов, скажем вспышки сверхновой или извержения вулкана, кипение воды в чайнике – преднамеренное событие; оно соотносится с ментальными состояниями агента, его внутренним осознанным содержанием, намерениями и убеждениями. Кто-

то захотел чая, и ему потребовалась горячая вода.

Можно, конечно, предположить, что действия отличаются в другом: что они состоят из движений тела живых существ или движений предметов, которые пространственно соединены с телом. Но искать контрпример долго не придется. Скажем, рефлекторные движения рук или ног явно нельзя отнести к действиям. Вздрагивания, кашель и храп не являются действиями, как и пробуждение или спотыкание [Davidson 1980, 44]. А полет радиоуправляемого самолета, скорее всего, – это действие, хотя сам пилот остается на земле. Убийство с помощью выстрела тоже, скорее всего, следует считать действием, хотя движение пули, непосредственно попадающей в жертву, не является движением тела агента.

Действия от прочих процессов также отличаются возможностью их телеологического объяснения. Они предполагают направленность событий в соответствии с намерениями, желаниями и убеждениями субъектов. Так, в случае с убийством направленность может соотноситься с желанием отомстить, войти в историю или получить коммерческую выгоду. Соотношение желаний, убеждений и действий агента аналитический философ Дж. Ким выражает с помощью следующего принципа:

(A) Если агент S желает что-то и убежден, что действие A обеспечит ему реализацию желания

оптимальным образом, S будет делать A².

В формулировке Кима неизбежной оказалась ссылка на оптимальность. Без нее прогноз включал бы в себя бесчисленное множество возможных действий, ведь одного и того же результата можно добиться разными способами. Сузить список помогает понятие оптимального действия – это реализация намерения с наименьшим количеством усилий в кратчайшие сроки и в соответствии с навыками агента. Но в данном случае эта оптимальность относительная. Для того чтобы сделать выбор образа действий, достаточно, чтобы сам агент считал способ действия оптимальным, не видел лучших альтернатив. При этом объективно могут существовать более оптимальные способы реализации желания.

Однако с одним этим разъяснением принцип все еще не будет эффективно работать. Он выглядит слишком упрощенным, ведь у агента редко бывает только одно желание. Нужно сделать поправку на то, что желание не входит в конфликт с другими желаниями и убеждениями. Скажем, желание выпить чашку горячего чая могло бы вступить в конфликт с другим, более сильным желанием – успеть на работу – и убеждением, что время выходить уже наступило. Тогда вода бы не закипела в чайнике, и он остался бы пустым и холодным. На применение обозначенного принципа накладываются и другие ограничения. Вот вы захотели выпить чаш-

² В таком виде этот принцип приводит Дж. Ким в книге «Философия сознания» [Kim 2006, 126].

ку чая, идете на кухню, а дверь в кухню закрыта на ключ и ключа нет. Принцип работает только в рамках допустимых возможностей. Конечно, на кухню можно было бы попасть и через окно или даже выломав дверь. Но здесь, скорее всего, опять возникнет конфликт желаний. Вряд ли желание выпить чаю будет сильнее, чем желание сохранить собственную жизнь или имущество неприкосновенными. Поэтому принцип нужно усовершенствовать так:

(Б) Если агент S желает что-то и убежден, что действие A обеспечит ему реализацию желания оптимальным образом, и при этом реализация желания физически возможна и не противоречит другим его более сильным желаниям, S будет делать A.

Помимо физических ограничений и конфликтующих желаний, препятствовать действию A могут также привычки. S, например, захотел утром горячего чая, но в чашку плеснул по привычке кофе. Действие может не последовать, даже если нет препятствий, в том случае, если само желание не проявляется достаточно сильно. Или не произойти потому, что оно противоречит представлению агента о собственном долге. Это представление может складываться из морального чувства, обычаев и практики. Агент может взять на себя обязательства и, несмотря на наличие противоречащего желания, эти обязательства соблюдать. Подобные случаи можно было бы также внести в качестве исключений в указанный базовый принцип. Но чтобы избежать дополни-

тельных усложнений, следует просто принять, что в принципе (Б) подразумевается желание достаточно сильное, чтобы инициировать действие, а оптимальный способ реализации не противоречит нормативным представлениям агента. В таком виде принцип, видимо, корректно фиксирует корреляцию между ментальными состояниями (убеждениями и желаниями) и действиями. Но остается не до конца ясным, на чем основана эта корреляция. Как именно желания и убеждения связаны с действием?

3. Два типа возможных связей между ментальными состояниями и действиями

Из принципа (Б) мы видим, что желания и убеждения по направленности коррелируют с действиями. Они предвещают результат действий. Кроме того, ментальные содержания участвуют в объяснении, рационализации самого действия. Почему человек вскипятил воду? Потому что хотел выпить чаю. Почему он совершил убийство? Потому что он надеялся попасть в историю. Но тот факт, что объяснение действия предполагает указание на эти состояния, не дает нам однозначности в понимании подчинительного союза «потому что». Дж. Ким замечает, что можно выявить как минимум два значения этого союза и, соответственно, два типа отношений. Иногда подчинительный союз обозначает (1) кау-

зальную связь, то есть указывает на причину, как, например, в предложении: «Здание обрушилось потому, что случилось сильное землетрясение». Здесь союз связывает два различных события, одно из которых было достаточным и необходимым основанием свершения второго. И эта каузальная связь подпадает под один из общих законов. Второе значение не предполагает каузальной связи и не является частным случаем общего закона. Союз «потому что» в таком случае разъясняет первое событие посредством расшифровки его значения: «Она стала профессиональной художницей, потому что рисует картины и выставляет их на выставках» или «Он стал совершеннолетним, потому что ему исполнилось восемнадцать лет». В этих предложениях речь не идет о двух различных состояниях или событиях. Художник и человек, профессионально рисующий картины, – синонимичные понятия. Наступление совершеннолетия и достижение восемнадцати лет – тоже одно событие, и слово «потому» здесь просто разъясняет это событие. Разъяснение здесь фактически указывает на (2) логическую связь.

Современных философов, которые считают, что между действиями и ментальными состояниями, объясняющими их, существует каузальная связь, называют *каузалистами*. Их противников – *антикаузалистами* (*noncausalist*). Последние убеждены, что между действиями и объясняющими их ментальными состояниями не существует каузальной связи, хотя может признаваться какая-то другая связь, на-

пример логическая [Ginet 2002]. До публикации Дэвидсоном статьи «Действия, мотивы и причины» (1963) доминирующей позицией среди аналитических философов был антикаузализм. После публикации Дэвидсона ситуация радикально поменялась. О результате изменения философ, профессор Университета Флориды

А. Мили пишет: «Лидирующим подходом в отношении как понимания, что такое действие, так и объяснения или определения действий является подход, опирающийся на каузальность. То, что может называться *стандартным каузализмом*, выражается в сочетании двух тезисов: (1) событие может являться действием в зависимости от того, что служит его причиной; (2) характерными объяснениями действий являются каузальные объяснения» [Mele 2009, 554]. Таким образом, Дэвидсон стал основоположником современного каузализма. Он представил аргументы, показывающие, что мотивы, объясняющие действия, являются их причинами. Эти доводы будут рассмотрены в ближайшее время. Но сначала следует рассмотреть аргументы в пользу антикаузализма, считавшиеся наиболее правдоподобными до 1963 г. Любопытно, что Дэвидсон косвенным образом внес вклад в развитие и этой позиции.

4. Доводы против каузальной связи действий и ментальных состояний

А. Трудности с доказательством психофизических законов

Традиционной моделью каузальности является номологическая модель³. Ее принимают многие современные философы. Модель предполагает, что каждый случай причинной цепи представляет собой частный случай действия общих законов. То есть событие a является причиной события b в том случае, если событие a подпадает под тип событий F , а b – под тип G , и существует закон, связывающий события типа F и G . Нагревание воды в чайнике до определенной температуры, к примеру, является причиной ее кипения, так как существует общая закономерность парообразования в жидкости при определенных температурах и давлении, выраженная в законах термодинамики. Аналогичная ситуация должна быть и в нашем случае. Но получится ли найти подобную регулярность между ментальными событиями и действия-

³ К примеру, ее принимали К. Гемпель и Д. Дэвидсон. «Согласно как Гемпелю, так и Дэвидсону, каузальность – это вопрос номологической достаточности: C является причиной E , если и только если существует закон, который соединяет события, подобные C , с событиями, подобными E » [Psillos 2009, 149].

ми, обнаружить психофизический закон? Кажется, что такие законы обнаружить, проверить или фальсифицировать проблематично.

Обычно гипотеза о наличии регулярности проверяется путем выделения предполагаемого действующего фактора, многократного воспроизведения его и наблюдения результатов. Если исходное событие является причиной, то ожидаемый результат последует всегда. Но в ситуации с убеждениями и желаниями это сделать практически невозможно: не получается полностью воспроизвести исходные условия и изолировать ментальную причину. Если действие уже осуществилось, нельзя исключать, что его причиной могли быть иные желания или убеждения. А если результата не последовало в случае воспроизведения желания, вполне вероятно, что появилось другое конфликтующее желание или субъект передумал. Ментальные состояния (желания в том числе) можно всегда совместить с отсутствием действия или с совсем другим действием. Поэтому с ментальными состояниями невозможно проверить наличие регулярности. Вот, к примеру, некто бросился бежать из-за того, что испугался. Страх и желание выжить, возможно, были причиной бегства. Но разве обязательно, что, если этот некто испугается в следующий раз, он побежит опять? Вполне вероятно, в следующий раз он замрет на месте (чтобы его не обнаружили или потому, что потерял способность двигаться) или, наоборот, бросится в атаку на противника.

Тогда, может быть, наличие психофизических законов можно подтвердить на своем примере, с помощью интроспекции, внутреннего самонаблюдения? Ведь когда я что-то желаю, я это делаю всегда. Достоверность этого суждения, однако, тоже не очевидна. Возможно, желание присутствовало раньше и повлекло действие, но, возможно, по действию я сделал вывод о собственном желании, то есть имела место инверсия. В оценках собственных состояний всегда имеется возможность новой интерпретации. Путем интроспекции такие законы проверить тоже трудно.

Но, может, объективно идентифицировать ментальное состояние и время его возникновения удастся в будущем? Может, это временное ограничение, и оно будет преодолено с развитием науки, появлением инструментов для точного сканирования мыслей? К примеру, в будущем появится возможность «увидеть» и выделить «желание выпить горячий кофе» (или его нейрофизиологический коррелят) и проследить всю последующую каузальную цепочку вплоть до движения рук и разогревания чайника? Существует, однако, аргумент в пользу того, что психофизические законы вообще не могут существовать. Его предложил Д. Дэвидсон. Вывод, который получается в результате этого рассуждения, он называет выводом об «аномальности ментального» [Davidson 2006b].

В. Аргумент аномальности ментального

Отличительной чертой ментальных содержаний является рациональность и согласованность. Невозможно допустить, чтобы субъект был убежден в противоречащих вещах: например, в том, что Москва находится на расстоянии 700 км от Петербурга и на расстоянии 1000 км от Петербурга. Даже когда субъект делает противоречивые высказывания, мы находим им непротиворечивые интерпретации. Так, высказывание «Я хочу поехать в Москву и не хочу туда ехать» мы склонны объяснять как сменяющиеся во времени состояния или как желание увидеться с родственниками в Москве, но нежелание стоять в бесконечных московских пробках. Высказывание «Мне нравится и не нравится Херст» мы не будем интерпретировать как свидетельство наличия противоречащих убеждений, но как выражение того, что некоторые работы Херста ценителю современного искусства нравятся, а некоторые – нет. Итак, требования логической непротиворечивости и связности являются обязательными для ментальных содержаний. Но подобных ограничений не существует для физических событий. Одни физические события нельзя лишь с помощью одной только логики вычислить из других. Физические законы случайны, а не необходимы. Из этих различий, по мнению Дэвидсона, следует, что области ментального и физического не могут быть связаны между собой

строгими законами. А если они не связаны законами, следовательно, согласно номологической модели каузальности, они не могут стоять в каузальной связи.

Для доказательства допустим обратное: предположим, что мы установили, что состояние мозга X связано с убеждением Y , то есть выявили такой-то психофизический закон. Тогда мы могли бы устанавливать и воспроизводить каждое отдельное убеждение без учета его соотношения с другими желаниями и убеждениями, даже противоречащими ему. Но это невозможно. Вероятно, мы столкнулись с противоречием, а значит, психофизические законы невозможны. Может быть, это является угрозой и для стандартного каузализма?

Ранее я указывал, что альтернативой каузальной связи может быть концептуальное отношение. Для того чтобы опровергнуть каузализм, необходимо показать, что именно концептуальная связь присутствует в отношении ментальных состояний и производимых ими действий.

С. Аргумент логической связи

Можно попытаться показать, что ментальные содержания находятся в концептуальной связи с действиями, и таким образом продемонстрировать, что объяснения с помощью ментальных содержаний не указывают на два различных, следующих друг за другом события (причину и действие), а просто расшифровывают, рационализируют действие. В этом за-

ключается суть Аргумента логической связи (АЛС). Считается, что «это самый значимый аргумент против каузализма» [Stoecker 2010, 603]. Его отстаивали Г. Райл и Э. Энском⁴. Аргумент удобней продемонстрировать на примерах.

Человек издали машет руками. Если спросить, почему он это делает, он ответит: «Я размахиваю руками, чтобы привлечь внимание». Очевидно, что в этом случае движения руками и привлечение внимания – это не разные события, а одно и то же, только с разными описаниями. Второе описание является объяснением первого потому, что указывает на значение первого действия, а не на его предшествующие состояния. (Это описание из более широкой перспективы, схватывающее более общий паттерн поведения.) Аналогичным образом, желания и убеждения, фигурирующие в нашем случае, возможно, являются не предшествующими событиями, а значениями действий. И логически они неотделимы от самого действия, в то время как причина должна быть логически независимой от следствия. Вот другой пример.

Гость за столом просит вас передать соль. Вы протягиваете ему солонку. Почему вы это делаете? Очевидно, вы поняли просьбу и хотите передать гостю солонку. Но желание передать солонку, разве оно отделимо от действия? Нельзя сказать, что вы передаете солонку без желания это делать (ведь никто вас не заставляет передавать соль, и это действие – не

⁴ Подробно Аргумент разбирается в статье Дж. Оттена «Возвращение к жизни Аргумента логической связи» [Otten 1977].

автоматизм), но вряд ли можно выделить желание в отдельный процесс. Если бы это был отдельный процесс, были бы способы его отдельно зафиксировать. Но как узнать о наличии подобного желания в отрыве от действия? Представляется абсурдным искать какое-то особенное ментальное состояние, «тягу передать соль». Каким образом эта тяга могла бы проявляться, кроме как в передаче соли? Некоторым из тех, кто размышлял над подобными вопросами, ответ кажется очевидным. По их мнению, действие и желание действия логически связаны. И все же это далеко не окончательный вывод. Ключевым оппонентом этого аргумента является Д. Дэвидсон.

5. Доводы Д. Дэвидсона в пользу каузальной связи

В статье «Действия, мотивы и причины» Дэвидсон защищает положения, которые до него считались несовместимыми. Первым из них является утверждение, что объяснения действий с помощью ссылок на ментальные состояния являются каузальными объяснениями. Вторым положением является утверждение о том, что у ментальных состояний, объясняющих действия, есть рационализирующая функция, то есть что ментальные состояния раскрывают мотивы действий. Для целей доказательства объясняющие ментальные состояния Дэвидсон называет «основными мотива-

ми» (*primary reasons*). Философ утверждает: «1. Для того чтобы понимать, как какой-либо мотив рационализует действие, необходимо и достаточно, чтобы мы понимали (как минимум в самом общем виде), из чего состоит основной мотив. 2. Основной мотив действия является его причиной» [Davidson 1980, 4]. Составом основных мотивов, или объясняющими ментальными состояниями, автор называет пару из а) стремления или склонности и б) соответствующего убеждения. Примером такой пары в случае действия «кипячение воды» служит желание выпить чаю и убеждение, что кипячение воды позволит приготовить чай. Дэвидсон последовательно защищает оба своих утверждения.

Первое утверждение Дэвидсон обосновывает с помощью анализа серии примеров. Так, он прибегает к следующему примеру: я переключил выключатель, включил свет и осветил комнату, не зная, что тем самым информировал грабителя о том, что я дома. Философ замечает, что включение света – это действие, но информирование грабителя – нет. У действия есть способ рационализации с помощью основного мотива, желания осветить комнату, и убеждения, что переключение выключателя приведет к этому. Информирование грабителя не имеет такой рационализации. Дэвидсон замечает, что иногда для рационализации используется только мотивирующее желание. Он считает, что в таких случаях мотивирующее убеждение просто подразумевается. Знание основного мотива рационализует действие, так как оно по-

казывает его телеологическую направленность. «В свете основного мотива действие предстает соответствующим определенным качествам агента (долгосрочным или краткосрочным, характерным для него или нет), и агент представляется в качестве Рационального Животного» [Davidson 1980, 8]. Но на каком основании автор считает, что основные мотивы являются причинами действий?

Ссылка на основной мотив является способом объяснения действия. Это объяснение рационализирует действие, показывая, что оно соответствует интенциям агента. Но это не означает, что объяснение с помощью ссылки на основной мотив является только рационализацией. Более того, есть основание считать, что объяснение с помощью ссылки на основной мотив является также каузальным объяснением. Таким основанием служит гипотетическая возможность случаев, в которых есть основной мотив, рационализирующий действие, есть действие, и тем не менее это действие выполняется по другой причине. «Сущностью отношения между действием и основным мотивом, который его объясняет, служит то, что основной мотив *обосновывает* поступок для агента» [Davidson 1980, 9].

Для примера возьмем случай, когда человек А случайно получает перевод средств, предназначавшийся человеку Б. А сообщает банковским служащим о том, что произошла ошибка [Lepore, Ludwig 2006, 4]. Основным мотивом является желание вернуть деньги тому, кому они предназна-

чались. Однако может быть такая ситуация, что существует еще один мотив: желание не попасть под налоговую проверку, которая обнаружит незаконное поступление средств. Оба желания, или обе пары желаний и убеждений, могут быть рационализацией действия А, но только один из мотивов может быть причиной. Именно этот мотив будет основанием поступка. Так Дэвидсон обосновывает второе положение. Для обоснования стандартного каузализма, впрочем, необходимо ответить на критику с помощью уже обозначенных аргументов. Во-первых, нужно показать, в чем заключается слабое место Аргумента логической связи (АЛС). И во-вторых, необходимо продемонстрировать, как соотносится убеждение о ментальной каузальности с собственными выводами Дэвидсона об отсутствии психофизических законов.

Дэвидсон предлагает несколько ответов на АЛС. Во-первых, он указывает на то, что, например, истинность высказывания «Я хотел передать соль» вовсе не влечет за собой истинность высказывания «Я передал соль». Я мог хотеть, но не передать. Между этими событиями нет необходимого следования. Во-вторых, что более важно, событие, которое определяет истинность высказывания «Я передал соль», не является интенциональным объектом в высказывании «Я хотел передать соль». Ведь если я передал соль, то сделал это каким-то определенным образом, в какой-то конкретный момент. А желание могло быть реализовано почти бесконечным числом конкретных физических действий. Я мог

бы передать соль сидя или стоя, правой или левой рукой, из рук в руки или бросив солонку гостю через весь стол. Учитывая такие различия желаний и реализующих действий, оправданно утверждать их логическую независимость и допустимость использования первых для объяснения вторых [Davidson 2006a, 25].

Дэвидсон считает, что логическая зависимость не существует между желаниями и действиями, но может существовать между некоторыми их описаниями. Как, например, в случае с высказыванием «причина пожара стала причиной пожара». Но всегда можно дать другие описания. Можно найти описания ментального без логической связи с действием. Как, например, в предложении «разожженный костер стал причиной пожара». Стандартной языковой практикой является описание события с помощью указания на причину. Но «описывать событие в терминах его причины – не то же самое, что объединять событие и его причину, и объяснение с помощью такого описания не исключает каузальное объяснение» [Davidson 1980, 14]. То есть можно «получение травмы» описать как «лучевой ожог». Но, по мнению Дэвидсона, из практики подобных описаний не следует то, что они указывают на одно и то же событие. Таким же образом, по его мнению, из отсутствия психофизических законов не следует, что между основными мотивами и действиями нет каузальной связи.

Дэвидсон, как я уже писал, склоняется к тому, что кау-

зальная связь должна быть частным случаем общего закона. Он также обращает внимание, что каузальная связь между мотивами и действиями в силу специфики мотивов действительно не может быть частным случаем какого-то закона. Но он предполагает, что мотивы могут быть описаны с помощью каких-то других дескрипций, и эти дескрипции могут быть описаниями частного случая закона. «Законы, существование которых необходимо, если мотивы являются причинами действий, не должны... быть сформулированы в терминах рационализаций. Из того, что причины класса событий (действий) попадают в определенный класс (мотивов) и существует закон, объясняющий каждое высказывание о каузальности, не следует, что существует закон, соединяющий события, классифицированные как мотивы, с событиями, классифицированными как действия, – классификацию можно провести и в нейрофизиологических, химических или физических терминах» [Davidson 1980, 17].

В качестве аналогии Дэвидсон предлагает рассмотреть метеорологический прогноз. Стандартные описания погоды включают в себя очень общие описания: «завтра во второй половине дня будут осадки, порывистый ветер, переменная облачность». Эти описания не могут фигурировать в законах метеорологии, так как не являются достаточно точными и построены на других концепциях. Однако это общее описание может быть переформулировано с большей точностью и с помощью других концепций, и тогда оно будет частным

случаем общего закона. Таким образом, Дэвидсон снимает противоречие психофизической аномальности представлениям номологической модели и обосновывает ментальную каузальность. Впрочем, есть и другие аргументы в ее пользу.

Без ментальной каузальности невозможны объяснение и рационализация многих социальных практик, в частности применение моральной ответственности. Люди часто считают других людей виновными в каких-то действиях или, наоборот, благодарят их за действия. Они сердятся, возмущаются или испытывают благодарность в отношении других. За свои поступки люди также испытывают гордость или чувство вины. Эти реакции определяются в современной аналитической философии как проявление моральной ответственности и результат признания свободы воли. Однако эти реакции были бы неуместными, если бы ментальная каузальность отсутствовала. Современный метафизик Дж. Ким пишет, что практика применения моральных категорий «очевидным образом предполагает, что наши ментальные состояния имеют каузальный эффект в физическом мире» [Kim 2008, 9]. Иначе на каком основании требовать от субъектов соответствия норме поведения, если поведение не связано с их ментальными состояниями? На каком основании благодарить или требовать ответственности, если желания и убеждения не играют каузальной роли? Как может быть свободен агент, если свобода проявляется в действии, но действие каузально не определено его ментальными состояниями: жела-

ниями и убеждениями? И как агент может нести ответственность за действия, если он является пассивным наблюдателем по отношению к ним?

Да и само понятие агента, разве не станет оно пустым без ментальной каузальности, ведь агент – это тот, кто имеет способность поступать в соответствии с собственным разумом. Если нет агента, нет и действия, ведь, как мы заметили в самом начале, связь с агентом и его ментальными состояниями отличает действия от всех прочих событий. Отрицая ментальную каузальность, возможно, придется вообще отрицать разум и ментальное, ведь само существование чего-то трудно представить без наличия у него каузальной силы. Это утверждается в так называемом Принципе элеатов. Согласно этому принципу, «нечто является реально существующим только в случае, если оно может быть участником каузального процесса»⁵. Поэтому вопрос о ментальной каузальности представляется в большей степени не эмпирическим вопросом, существует она или нет, а философским: как возможно ее существование.

Относительно важности ментальной каузальности категорически высказался другой современный аналитический философ, Дж. Фодор: «Я не полностью убежден, насколько важно, что ментальное – это физическое; и еще меньше в том, что важно, чтобы мы могли это доказать. Однако если неверно то, что мое желание каузально отвечает за

⁵ Принцип сформулирован по статье [Colyvan 1998].

мое действие, щекотка за чесотку, мое убеждение за мои выражения... если все это неверно, тогда все, в чем я убежден, – ложь, и это конец света» [Цит. по: Kim 2006, 181]. Фодор справедливо заметил, что от ментальной каузальности зависит истинность человеческого знания. Ведь переход от восприятия к убеждению и от одного убеждения к другому невозможно помыслить без каузальных связей между ментальными состояниями. С Фодором согласны и большинство российских философов. Так, по данным опроса Московского центра исследования сознания, 92 % считают сознание каузально эффективным [Беседин, Васильев, Волков, Кузнецов 2017]. С этим согласен и я. Однако это согласие ведет к проблеме сочетания ментальной каузальности и физикализма. В современной философии эта проблема выявляется в связи с *Аргументом исключения*.

II. Проблема ментальной каузальности в современной аналитической философии

1. Аргумент исключения Дж. Кима против ментальной каузальности

Аргумент исключения, судя по количеству комментариев и ответных статей, – один из наиболее значимых аргументов современной аналитической философии в области ментальной каузальности. Первые версии его были разработаны Ч. Тэйлором [Taylor 1964] и Н. Малкольмом [Malcolm 1968]. В конце 80-х он был актуализирован американским метафизиком Дж. Кимом, а затем подробно разработан им же в статье «“Нисходящая каузальность” в эмерджентизме и нередуктивном физикализме» в 1992 г. Как пишет А. Кузнецов в своей диссертации, «Ким – философ, вокруг которого строятся дискуссии о проблеме ментальной каузальности» [Кузнецов 2016b, 52]. В наиболее поздней монографии «Физикализм или что-то около того» Ким представляет обновленную версию Аргумента, вписанную в более широкую теорию отношения ментального и физического [Kim 2008, 13].

Рассуждение Кима основано на совмещении нескольких

вполне правдоподобных принципов. Это три утверждения: Принцип каузальной исключительности (ПКИ), Принцип каузальной замкнутости физического мира (ПКЗ) и Принцип отсутствия систематической сверхдетерминации (ПОСС). Принцип каузальной исключительности фиксирует, что в стандартной ситуации причина события может быть только одна:

(ПКИ) Если событие e имеет достаточную причину c в момент t , никакое событие v в t , отличающееся от c , не может быть причиной e (кроме случаев сверхдетерминации).

Согласно этому принципу утверждается, что достаточная причина у события должна быть одна, кроме специальных случаев сверхдетерминации. Под случаями сверхдетерминации имеются в виду случаи, когда событие могло произойти по любой из нескольких не связанных между собой причин. К примеру, смерть от расстрела стрелковым подразделением – это сверхдетерминированное событие, так как попадание каждой пули в жизненно важные органы является смертельным и пули двигаются независимо друг от друга. Принцип каузальной исключительности – для многих это установка по умолчанию. Она может даже казаться аналитической истиной. Большая часть исследователей среди *relata* причинно-следственной связи насчитывают одну достаточную причину и одно следствие. Если причина достаточная, то она будет одна.

Второй важный принцип – это Принцип каузальной замкнутости физического мира (ПКЗ). На этом принципе базируется научное знание, в частности понятие физического закона. И несмотря на то, что существуют отдельные любопытные попытки оспорить этот принцип, он все-таки представляет правдоподобное положение. Согласно ПКЗ:

(ПКЗ) Каждое физическое событие в момент t имеет в качестве своей достаточной причины другое физическое событие в момент t .

Ссылка на конкретную отметку по временной шкале нужна Киму для того, чтобы исключить конфликт причин, разнесенных во времени. Одно и то же событие может быть результатом нескольких следующих друг за другом событий. В этом виде, согласно ПКЗ, все процессы и действия – вспышка сверхновой, извержение вулкана, взлет беспилотника и кипение воды в чайнике – в равной степени обусловлены предшествующими физическими событиями. Чтобы объяснить любые события, достаточно только физических причин, в природе нет места чудесам.

Впрочем, этот принцип можно немного адаптировать, ведь кто-то может допускать беспричинные события или случайности. Вместо того чтобы спорить по этому поводу, здесь можно просто несколько ослабить этот принцип, представив его в таком виде:

(ПКЗ¹) Если физическое событие имеет причину в

момент t , оно имеет достаточную физическую причину в момент t .

В таком виде этот принцип допускает беспричинные события (как и предыдущий принцип), утверждает лишь наличие физических причин у событий с причинами. И он не должен вызывать сомнения.

И наконец, последний интересующий нас принцип – Принцип отсутствия систематической сверхдетерминации:

(ПОСС) Не существует систематической сверхдетерминации физических событий.

Я уже разъяснил, что имеется в виду под сверхдетерминацией, и привел пример, когда она имеет место. Но подчеркну, что сверхдетерминация – это особенный случай. Систематическая, регулярная сверхдетерминация представляется невероятной. Одно дело – допускать единичный случай наступления смерти в результате расстрела стрелковым подразделением, а другое – утверждать, что всегда, когда кто-то погибает от пуль, стреляющих как минимум двое. Второе кажется невероятным.

Итак, что получается из соединения этих принципов? По мнению Кима, ПКЗ, ПОСС и ПКИ в сочетании образуют серьезную трудность для интерпретации ментальной каузальности при условии, что ментальные состояния не отождествляются с физическими процессами. Ход рассуждений философа таков. Допустим, ментальное событие M является при-

чиной физического события, действия P^* . Обозначим это $M \rightarrow P^*$. При этом, согласно ПКЗ, у события P^* есть также достаточная причина P – другое физическое событие, некоторое состояние мозга, одновременное с событием M . То есть $P \rightarrow P^*$. Если $M \neq P$, то у P^* две причины: ментальная и физическая. То есть P – сверхдетерминированное событие. Иными словами, каждое из этих предшествующих событий достаточно для реализации события P . Но, согласно ПКЗ, физическая причина, предшествующее состояние мозга, обязательна, а ментальная – нет: согласно условию, мы имеем дело с физическим событием, у которого есть причина, а значит, у него должна быть достаточная физическая причина (согласно ПКЗ), тогда как наличие у него ментальной причины не обязательно. Значит, событие P^* произошло бы, даже если бы M отсутствовало. Следовательно, ментальные события не вносят никаких изменений в ход дел. Они *эпифеноменальны*, то есть не имеют каузальной силы. События развивались бы точно так же, даже если бы ментальных состояний не было. «Очевидный вывод этого аргумента в том, что каузальное действие ментального на физическое является иллюзией; оно никогда не происходит. Это эпифеноменализм, как минимум в отношении физических событий. Это не исключает влияния одних ментальных событий на другие ментальные события. Но если ментальные события, как убеждения и интенции, никогда каузально не влияют на движения те-

ла, это лишает агента возможности существования...» [Kim 2006, 217].

Конечно, чтобы попытаться спасти ментальную каузальность, можно допустить тождество M и P : $M = P$. Но это жертва очень значительная. На первый взгляд, мы спасем ментальную каузальность. Но при этом придется отрицать особый статус ментального и его специфические свойства, в том числе возможность множественной реализации. А это – тоже общепризнанный тезис. Ментальные свойства могут быть реализованы в разных случаях и у разных индивидов разными физическими состояниями. К примеру, боль как ментальное свойство может быть реализована возбуждением различных нервных окончаний и активацией разных нейронных сетей. Таким образом, Аргумент исключения фактически ставит перед нами дилемму: либо ментальная каузальность невозможна и ментальные состояния в лучшем случае эпифеноменальны, либо они тождественны физическим процессам, а значит, специфических ментальных свойств не существует. Обе альтернативы крайне нежелательны. Эта дилемма и представляет суть проблемы ментальной каузальности.

2. Аргумент супервентности против каузального влияния одних ментальных состояний на другие

Впрочем, в приведенных рассуждениях все еще присутствует открытый ход. Выше был рассмотрен случай каузального влияния ментального на физическое. Аргумент исключения приводит к выводу, что это влияние невозможно. Но может быть, ментальное хотя бы способно влиять на ментальное, то есть одни ментальные состояния могут быть причиной других? Если одни ментальные состояния влияют на другие, справедливым будет, по крайней мере, сказать, что убеждения могут основываться на воспоминаниях, желания – на убеждениях, а выводы – на умозаклучениях. Это хотя бы отчасти даст основания к утверждению существования и действенной силы разума.

К сожалению, и здесь нас ожидает препятствие – еще один аргумент Кима – Аргумент супервентности (АС). Он строится на уже знакомых принципах: ПКЗ, модификации ПКИ и Тезисе о супервентности (ТС). Тезис о супервентности представляет собой ядро физикалистских позиций, он разделяется практически всеми, кто придерживается доминирующих сегодня философских взглядов. С ним согласны сторонники редукционистского физикализма (например, последователи теории тождества), функционалисты и даже эмерджент-

тисты. В основе ТС лежит представление о том, что полностью идентичные в физическом смысле системы будут полностью идентичны и в плане ментальных состояний. ТС указывает на то, что ментальное не является независимым и не существует автономно от физического, а, наоборот, *супервентно*, то есть зависимо и определено базовыми физическими свойствами. ТС имеет множество формулировок, к некоторым из которых мы обратимся ниже, однако для целей своего аргумента Ким представляет его в самом общем виде:

(ТС) *Все, что происходит в ментальной жизни, полностью определено физическими процессами.*

Вот что получается, когда мы анализируем каузальное влияние ментального на ментальное с учетом Тезиса о супервентности. Предположим, ментальное событие M является причиной другого ментального события, M^* . Согласно ТС, однако, ментальное событие M^* в качестве своей базы должно иметь некое физическое событие P^* , ведь ментальное не существует автономно. Теперь применим ПКИ, точнее вариант ПКИ. В исходном виде этот принцип не совсем подходит, ведь он относится к случаям каузальной связи. Но его можно обобщить до любых генерирующих (определяющих) связей, в том числе и отношения супервентности. Тогда он будет звучать так:

(ПКСИ) *Если событие e определено событием c – каузально или отношением супервентности, –*

тогда событие e не может быть определено другим, отличающимся от c событием, кроме случаев сверхдетерминации.

Применив этот принцип, получим: если событие M^* имеет достаточное основание в событии c , это и есть его единственное основание (кроме случаев сверхдетерминации). Но которое из двух перечисленных оснований, M или P^* , является достаточным? M , вероятно, детерминирует M^* горизонтально, с течением времени, а P^* – вертикально, как база для супервентного свойства. Известно, что, случись P^* , M^* произошло бы с необходимостью. Тогда достаточной базой является P^* . То, что мы наблюдаем в этом случае, сам Ким называет максимой Эдвардса (в честь американского философа и теолога Джонатана Эдвардса): «Существует конфликт между вертикальной детерминацией и горизонтальной каузальностью. Фактически вертикальная детерминация исключает горизонтальную каузальность» [Kim 2008, 36]. Тогда остается лишь один вариант влияния ментального на ментальное – M может стать причиной M^* , став предварительно причиной базового события P^* . Единственный способ каузально влиять на супервентное свойство – только через каузальное влияние на его базу. Но это шах, ведущий к мату. Ведь это случай каузального влияния ментального на физическое. А в отношении возможности подобного влияния был уже вы-

несен отрицательный вердикт в связи с Аргументом исключения. Круг замкнулся, и теперь мы в полной мере наблюдаем трудность философской проблемы ментальной каузальности.

Похоже, что ментальные состояния не могут каузально влиять не только на физические состояния, но и на другие ментальные состояния. Кажется, общепринятые взгляды на ментальное и его действенность непоследовательны. И попытки совместить веру в ментальную каузальность с важными метафизическими и логическими принципами приводят нас к патовой ситуации. Дж. Ким так обобщает эту ситуацию: «Следует заключить тогда, что проблема сознание-тело для нас, желающих быть физикалистами, разделяется на две проблемы: проблему ментальной каузальности и проблему сознания, и они вместе представляют самый большой вызов для физикализма. Если физикализму суждено уцелеть в качестве нашего мировоззрения, он должен показать, каково наше место в физическом мире, и это означает, что он должен показать наш статус как сознательных существ, обладающих возможностью влиять на окружающую среду благодаря сознанию и ментальности. Аргументы, представленные выше, уже показали, что физикализму не удастся сохраниться без изменений...» [Kim 2008, 31].

За последние несколько десятилетий вопросу о соотношении ментального и физического (проблема «сознания») и проблеме ментальной каузальности были посвящены мно-

гие сотни работ. Значительная их часть написана философами. Но растет также количество исследований, предпринятых учеными-когнитивистами, нейрофизиологами, психологами, врачами, физиками. Философия, кажется, теряет эксклюзивные права в этой области. Но среди исследователей остается что-то вроде общего консенсуса: физикалистская позиция продолжает преобладать. Это позиция, согласно которой все в окружающем нас мире в конечном итоге состоит из физических объектов и зависит от физических свойств. С появлением новых книг растет количество аргументов, и авторы с критикой набрасываются на тексты своих коллег, так что у любопытствующей публики создается полное ощущение горячего спора. Но если в действительности спор и имеет место, этот спор скорее о деталях. Оригинальность авторов обычно выражается в способе аргументации, но не в явной новизне позиции. Одним из редких исключений являются тексты нашего коллеги, российского философа В. В. Васильева. Он предлагает провести ревизию физикализма.

III. Решение проблемы ментальной каузальности в локальном интеракционизме

1. Ключевые положения локального интеракционизма

По нескольким статьям и трем монографиям⁶, опубликованным в последнее время, можно проследить оригинальную философскую позицию, которую защищает российский философ В. В. Васильев. Автор называет ее «*локальным интеракционизмом*». В соответствии с этой позицией утверждается специфический, нередуцируемый характер ментального. С точки зрения автора, квалиа, качественные характеристики воспринимаемого во внутреннем опыте, не могут быть сведены к физическому, к состояниям мозга. При этом, по его представлениям, квалиа не просто сопутствуют физическим процессам. Они не эпифеноменальны, не являются «номологическими бездельниками», а ре-

⁶ См. работы Васильева В. В. «Трудная проблема сознания» [2009], «Сознание и вещи: Очерк феноменилистической онтологии» [2014], «В защиту классического компатибилизма: Эссе о свободе воли» [2017], а также [Васильев 2016] и др.

ально влияют на поведение. Без них мозг не мог бы работать так, как он работает, поэтому существование квалиа эволюционно оправданно. Но как тогда быть с Принципом каузальной замкнутости физического? Достоинством и уникальной особенностью локального интеракционизма является попытка совместить тезис о ментальной каузальности и некоторый вариант ПКЗ. Положения локального интеракционизма можно резюмировать в следующих тезисах:

Онтологический статус сознания: Ментальные состояния существуют как субъективные переживания и нередуцируемы к физическим.

Ментальная каузальность: Ментальные состояния каузально влияют на поведение.

Связь с состояниями мозга: Они не могут быть полностью обусловлены только процессами в мозге.

Каузальная замкнутость: Принцип каузальной замкнутости не работает в некоторых локальных объектах, но сохраняется глобально в мире.

Рассмотрению этой позиции целиком следует посвятить отдельное исследование. Я же попытаюсь в этой главе разобрать один из интереснейших аргументов, который изобретает автор «Трудной проблемы сознания» – Аргумент каузальных траекторий (АКТ). Это рассуждение было разработано Васильевым в 2004 г. при размышлении над книгой «Сознающий ум» Д. Чалмерса [Васильев 2009, 216]. На этом аргументе основывается важная часть теории локального ин-

теракционизма и решение проблемы ментальной каузальности.

2. Аргумент каузальных траекторий в пользу ментальной каузальности и против локальной супервентности

Аргумент сформулирован в виде трех посылок и вывода. Первая посылка – это утверждение о том, что одинаковые события могут иметь разные причины. Вот шар попал в лузу, он мог там оказаться самыми различными способами, в результате самых различных соударений и движения по множеству траекторий. Как, впрочем, любая вещь могла оказаться там, где она есть, множеством различных способов. Посылка выглядит бесспорной. Вторая посылка – это констатация того, что память отражает события, происходившие с субъектом в прошлом, то есть отражает каузальную историю субъекта. Конечно, память может представлять каузальную историю ошибочным образом, но это скорее исключение, чем норма [Васильев 2009, 217]. Столь же несомненно, по мнению российского философа, что воспоминания представляют собой приватные квалитативные состояния и образуют основу убеждений, желаний и других осознанных состояний субъекта. Третья посылка указывает на корреляцию поведения и желаний. Эта посылка не отрицается даже эпифеноменалистами. Эпифеноменалист, правда, считает, что

корреляция не имеет отношения к причинно-следственной цепи. А вот представленный аргумент заставляет нас склониться к обратному.

Объединяя посылки, Васильев получает 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... 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... скоррелированы с моим поведением (третья посылка), то, обладая другими убеждениями и желаниями, я вел бы себя иначе, чем веду себя в данный момент...» [Васильев 2009, 217–218].

Итак, из аргумента следует, что, находясь в том же самом физическом состоянии, агент мог бы вести себя различным образом. И эти различия обуславливаются только его качественными ментальными содержаниями, различными убеждениями и желаниями. И никакой анализ наличного физиче-

ского состояния не способен дать точный прогноз действий субъекта. Ментальные состояния действительны, а значит, эпифеноменализм ложен. Но опровержение эпифеноменализма, с моей точки зрения, не самое критичное из следствий Аргумента каузальных траекторий. Гораздо более значимым мне кажется опровержение того, что является общим местом современных теорий сознания, – Тезиса о супервентности ментальных содержаний на состояниях мозга (ТС).

Понятие супервентности обозначает зависимость одних свойств или фактов от других. Если свойства типа *A* супервентны на свойствах типа *B*, то два объекта не могут различаться в свойствах *A* без различий в свойствах *B*. В современный философский дискурс это понятие ввел уже упомянутый Д. Дэвидсон в знаменитой статье «Ментальные события» (1970): «Хотя согласно позиции, которую я занимаю, отрицается существование психофизических законов, она совместима с мнением, что ментальные характеристики в некотором смысле зависят, или супервентны на физических характеристиках. Здесь супервентность означает невозможность двух событий, не отличающихся в физическом смысле, но отличающихся в ментальном отношении, невозможность изменения объекта в ментальном отношении без изменений в физическом отношении» [Davidson 2006b, 111]. Несмотря на то что Дэвидсон защищает свой вариант решения проблемы сознания, тезис о супервентности ментального на физическом является распространенным. И его

принятие могло бы служить критерием минимального физикализма. Его разделяет большинство философов: дуалисты свойств, эмерджентисты, функционалисты, сторонники теории тождества, элиминативисты, ведь ТС – это не решение вопроса о сознании и ментальной каузальности, а современная его постановка: как объяснить ментальное (сознание), если оно супервентно на физическом? Тем важнее оказывается аргумент, который подвергает этот тезис сомнению, отрицает супервентность. Впрочем, здесь необходимо сделать некоторые уточнения.

Возможны как минимум три формулировки Тезиса о супервентности. (1) Ментальное супервентно на физическом в том случае, если верно, что если нечто X имеет ментальное свойство M , то имеется физическое свойство P такое, что X обладает P , и необходимо, что всякий объект, обладающий P , наделен M . (2) Ментальное супервентно на физическом в том смысле, что объекты, совершенно одинаковые в своих физических свойствах, не могут отличаться относительно ментальных свойств. (3) Ментальное супервентно на физическом в том смысле, что миры, одинаковые во всех физических отношениях, одинаковы и во всех ментальных отношениях; то есть миры, одинаковые физически, совершенно одинаковы в целом⁷. Сила этих утверждений различна. Наиболее сильным является утверждение (1), так как оно фиксирует супервентность даже одного изолированного свойства.

⁷ Варианты Тезиса о супервентности приведены по книге [Kim 2006, 9–10].

А наиболее слабым – (3), так как в нем утверждается супервентность только миров в целом. Если условия, описанные в утверждении (1), выполняются, то будут выполняться и условия (2). А если выполняются условия по утверждению (2), автоматически будут выполнены и условия по утверждению (3). Определение (3) фиксирует отношение глобальной супервентности, а определения (1) и (2) – локальной. Для примера возьмем свойство «быть банком». Это свойство не имеет локальной супервентности, но имеет глобальную. Так, здание может быть одним и тем же, но является ли это здание банком или, скажем, музеем, зависит от того, какой статус оно имеет и как используется обществом. Аналогично, свойство «быть оригинальной картиной Ван Гога» также не имеет локальной супервентности. Пусть 1000 картин не отличаются друг от друга ни на один атом, тем не менее одна из них может быть подлинной картиной Ван Гога, а другие – точными копиями. Это зависит не только от самой картины, но и от внешних по отношению к ней факторов. Итак, как соотносятся 3 варианта Тезиса о супервентности с выводом Аргумента каузальных траекторий?

Очевидно, реальная возможность существования двух идентичных людей с разными ментальными состояниями, допускаемая в качестве результата этого аргумента, противоречит варианту (1) ТС. Но сам Васильев сообщает, что его выводы не противоречат «глобальной супервентности»: «последняя формулировка выражает так называемую “глобаль-

ную 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» [Васильев 2009, 243]. Что же получается? Перед нами – неожиданная и удивительно красивая картина мира. Ментальные состояния отличаются от физических. При этом они не эпифеноменальны, а действенны. Они влияют на поведение. Но это влияние не разрушает законы природы, в том числе фундаментальный закон сохранения энергии.

Ментальное является посредником, условием влияния нелокальных физических факторов, а локально – причиной поведения. Ментальная каузальность защищена, тщетность усилий ученых найти корреляты ментальным состояниям объяснена, ПКЗ сохранен, и... в качестве бонуса получен ключ для решения двух философских проблем: «трудной проблемы сознания» и проблемы тождества личности. Из подобной картины следует, что сознательные состояния существуют потому, что необходимы для генерации человеческого поведения (без них поведение разумного человек было бы просто невозможно). И что тождество личности не связано только лишь с тождеством тела, но также и с тождеством каузальной истории, зафиксированной в ментальных состояниях. Что ж, если это построение, помимо своих бесспорных эстетических качеств, сможет выдержать критический анализ, то это крупный успех и даже прорыв в философии.

Итак, учитывая значимость выводов, вернемся к аргументу еще раз и попытаемся для начала проверить, действительно ли он опровергает Тезис о локальной супервентности, сохраняя при этом глобальный вариант. Именно в этом случае он совместим с ПКЗ, вписывается в современную научную картину мира и согласуется с базовыми естественными установками.

А. Локальная или глобальная супервентность?

Вполне вероятно, что в исходном виде аргумент действительно направлен исключительно против локальной супервентности ментального на физическом. Но есть вероятность, что, если аргумент верен, его вариацию можно использовать также и против глобальной супервентности. Тогда рассуждения строились бы следующим образом. (1) У одинаковых событий могут быть разные причины. Таким образом, если текущее состояние мира представляет собой гигантский комплекс событий, эти же события могли произойти в ходе других причин. (2) Память людей обычно отражает события, которые происходили с ними в прошлом. Но если бы у них было другое прошлое, их память имела бы другое содержание. (В) Следовательно, возможно, что существовал бы мир, полностью идентичный нашему, но отличающийся от нашего содержанием ментальных состояний населяющих его лю-

дей. Это следствие противоречит Тезису о глобальной супервентности – тезису о том, что миры, одинаковые физически, совершенно одинаковы в целом. Выходит, что Аргумент каузальных траекторий направлен и против глобальной супервентности?

Васильев предполагает подобного рода замечания и заведомо не согласен с выводом. Скорее всего, он критиковал бы первую посылку в нашем изложении. Хотя одно сингулярное событие (используя терминологию автора) может иметь множество предшествующих событий в качестве причины, весь комплекс событий, составляющий состояния мира в момент t , мог бы иметь только одно предшествующее состояние – то, которое было в действительности. То есть у текущего совокупного события, *События А*, может быть только один antecedent, *Событие В* (*ТСС – Тезис о совокупном событии*). Но каковы основания считать именно так?

У сторонника локального интеракционизма есть две тактики защиты этого Тезиса. Одна основывается на анализе принципов здравого смысла. По мнению Васильева, эти положения не могут быть оспорены, так как отражают устройство наших познавательных способностей. Одним из базовых положений Васильев считает убеждение «в соответствии прошлого и будущего» [Васильев 2009, 206]. На нем построены другие: вера в существование причин у каждого события, убеждение в том, 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» [Васильев 2009, 208], Принцип каузальной замкнутости физического. По мнению автора, аксиома соответствия прошлого и будущего обуславливает также веру в возможность существования только одной причины, только одного События А, предшествующего Событию В. Мне кажется, что следование здесь далеко не очевидно. Но в данном случае я не берусь оценивать умозаключения Васильева, так как метод философствования, который применяет здесь автор, частично выходит за рамки аналитического дискурса и представляет собой вариант аналитической феноменологии, довольно редко использующийся современными философами. Впрочем, к выводам автора проложен и другой путь, который проще проследить и проверить. Он представлен в другой книге Васильева, «Сознание и вещи» [Васильев 2014].

Философ сравнивает два описания события: отскок мяча от поверхности и отскок мяча от поверхности на такое-то расстояние. Оба события могут быть следствиями множества причин. Например, причиной могли быть удар по мячу другим мячом, или теннисной ракеткой, или неустойчивое исходное положение мяча на плоскости. Но второе описание является более конкретным и в этом смысле сужает горизонт возможных причин (предшествующих состояний). Кажется, что конкретизация описания событий уменьшает возможные antecedentes события. Отскок мяча на двести метров мог иметь гораздо меньше причин, чем просто от-

скок мяча. А состояние мозга в момент t является еще более «конкретным событием», чем отскок мяча на двести метров, и потому может иметь еще меньшее количество причин. Но явно не одну. Только предельно конкретное событие имеет исключительно одну причину. Таким событием и является Событие, или совокупное содержание всего чувственного опыта в момент t , – считает автор.

Кажется, эти рассуждения не являются полностью безупречными. Они строятся на представлении о снижении количества возможных причин при конкретизации описания. Но разве они убеждают нас в том, что совокупное событие⁸ может иметь только одну причину? Возьмем сингулярное событие и создадим для него исчерпывающее конкретное описание. Это описание будет бесконечно длинным, пусть так. Но будет ли у этого конкретного события возможность в качестве антецедента иметь больше одной причины? Кажется, что да, по крайней мере, в соответствии с посылкой 1 исходного аргумента. Вне зависимости от того, сколь точным, конкретным будет описание одного сингулярного события, к нему может вести больше одной каузальной траектории. А теперь подумаем о Событии. Разве Событие, совокупное событие, текущее состояние мира, по определению не является совокупностью всех конкретных сингулярных событий? Если так, то его предшественниками могут быть варианты ан-

⁸ Васильев использует термин примерно в том же смысле, «что Витгенштейн понимал под миром» [Васильев 2009, 241].

тецедентов всех конкретных сингулярных событий. И таких предшествующих состояний должно быть больше чем одно. Я даже думаю, что чем больше сингулярных конкретных событий входит в совокупность, тем больше вариаций ей может предшествовать.

К примеру, пусть комбинация шаров на бильярдном столе такова, что игрок может забить шар в лузу только с отскоком от правого или левого борта. Если шар уже попал в лузу, и мы не видели, как он был забит, мы должны допустить, что он туда попал одним из двух вариантов. Теперь представим, что футбольное поле или даже пустыня Сахара заставлена бильярдными столами с аналогичным расположением шаров, и у каждого стола находится по одному игроку. Вот все шары оказались в лузах. Какое количество комбинаций могло предшествовать попаданию всех шаров в лузы? Очевидно, что это число больше, чем число комбинаций, предшествующих попаданию одного шара. Мне кажется, таким же образом дело должно обстоять и с совокупным событием. А это значит, что, если Аргумент каузальных траекторий верен, он, вероятно, опровергает не только локальную супервентность, но и глобальную, то есть все три варианта формулировки ТС.

В таком случае ментальное становится полностью автономным от физического, парящим свободно над физическим миром. А это – нарушение ПКЗ, одного из принципов здравого смысла. Теория, которая ведет к таким следствиям,

вряд ли может выстоять. Впрочем, не стоит торопиться. Даже если нам удалось опровергнуть второй путь защиты ТСС, надежным мог быть первый. А для анализа его у нас недостаточно инструментов. Думаю, в таком случае оправданно вернуться к исходному аргументу и продумать его посылки. На мой взгляд, достаточно будет ограничиться первыми двумя. Ведь для получения интересующего нас вывода (о ложности ТС) достаточно только их. Мы сократим аргумент и приведем его в несколько адаптированный вид.

В. Разъяснение двусмысленности в аргументе

(1) Реально возможно, чтобы мой мозг пришел к состоянию C в результате различных событий (A или B и т. д.).

(2) Реально возможно, чтобы память отражала события, происходившие с субъектом в прошлом.

(В) Реально возможно, чтобы мой мозг находился в некотором состоянии C , но моя память при этом могла бы отражать одно из различных предшествующих событий (A или B и т. д.).

В новой формулировке аргумента сделаны некоторые правки. Первая посылка конкретизирована в отношении возможных каузальных состояний мозга, ведь именно об этом в ней речь. Введено уточнение: каждая из посылок

представлена не как факт, а как реально возможный факт. Это кажется уместным дополнением, так как только в этом случае послылки становятся верными. Мозг может прийти к некоторому состоянию C разными путями, хотя и не обязательно к нему придет. А память реально может отражать прошлые события, хотя и неверно, что она отражает все прошлые события. Вывод сокращенного аргумента опровергает ТС (1). И, насколько я понимаю, с такой формулировкой должен быть согласен автор Аргумента каузальных траекторий.

Однако в таком виде в аргументе проявляется некоторая двусмысленность. Реально возможно, чтобы мозг пришел к состоянию C в результате разных событий, реально возможно, чтобы память отражала произошедшие с субъектом события. Но в этих двух послылках – можно возразить – речь идет совсем о разных событиях! В одном случае – о событиях, происходящих в нейронных сетях, а в другом – о биографических событиях. Память не отражает каузальных цепей нейронных событий, даже память нейрофизиолога-мнемониста, это реально невозможно. Значит, связь между послылками отсутствует? И из посылок вывода о ложности ТС вовсе не следует?

Конечно, дело обстоит не так. Двойственность из аргумента устранить возможно. Но тогда, мне представляется, рассуждение должно иметь более сложный характер. Назовем это рассуждение АКТ 1:

(1) Реально возможно, чтобы мой мозг пришел к состоянию C в результате различных нейронных событий (A , B и т. д.).

(2) Другие нейронные события могли соответствовать другим биографическим событиям.

(3) Реально возможно, чтобы память отражала биографические события.

(В) Реально возможно, чтобы мой мозг был в одном и том же состоянии C , но моя память отражала различные биографические события (A , B и т. д.).

Посылки представляются истинными. И следовательно, истинным может быть вывод. Правда, есть существенное условие. Для истинности вывода нужно не только, чтобы реально возможными были каждая из описанных в посылке ситуаций, но и их конъюнкция: все события должны быть возможными одновременно. Ведь реально возможно начать игру в шахматы и с хода пешкой, и с хода конем, но нельзя начать игру сразу с двух ходов, конем и пешкой. Мне кажется, что совпадение ситуаций, описанных в посылках аргумента, все-таки невозможно. Невозможна ситуация, когда есть два двойника, и их мозги находятся в идентичных состояниях, но пришли к этим состояниям в результате разных нейронных событий, а также разных автобиографических событий, и эти двойники имеют разные воспоминания.

3. Возражения на аргумент каузальных траекторий

А. Аргумент-двойник

Доводом в пользу невозможности этой ситуации служит аналогичный аргумент, Аргумент-двойник, который хотя и имеет, по моему мнению, ключевые черты АКТ 1, приводит к очевидно ложному выводу. Если Аргумент-двойник не работает, но принципиально не отличается от АКТ 1, то и АКТ 1 не работает. Я предлагаю рассмотреть рассуждение, которое я назову АКТ 2:

(1) Реально возможно, чтобы жесткий диск моего компьютера пришел в текущее состояние различными путями.

(2) Файловая система хранит информацию (компьютерная память) о последовательности изменений файлов на диске.

(В) Реально возможно, чтобы жесткий диск моего компьютера был в текущем состоянии, но файловая система хранила другую информацию о последовательности записи файлов на диск.

Это рассуждение похоже на АКТ 1, и доводы, приводимые в пользу истинности посылок АКТ 1, полностью применимы и в новом аргументе. Диск компьютера – такая же

физическая система, как мозг человека. Следовательно, текущему состоянию диска могли предшествовать различные состояния. Файловая система в качестве метаданных хранит информацию о времени создания, доступа и модификации каждого файла. Эта информация обычно представляется правильным образом и отражает индивидуальную каузальную траекторию жесткого диска. Отклонения, конечно, возможны, но мы говорим о норме. И это – общее представление, разделяемое как любителями, так и профессиональными программистами. И тем не менее вряд ли кто-либо станет утверждать отсутствие локальной супервентности данных файловой системы на физическом состоянии жесткого диска и допустит возможность ситуации, описанной в выводе АКТ 2. Но почему? Почему вывод АКТ 1 представляется правдоподобным в случае с памятью человека, а аналогичный вывод в случае с памятью компьютера в АКТ 2 – неправдоподобным? Может, нам стоит поэкспериментировать с жесткими дисками и экспериментально проверить работу Принципа супервентности в компьютерных системах?

И все-таки в случае с механическими системами в локальной супервентности информации на носителе вряд ли можно сомневаться. Здесь с нами согласится и Васильев. Об этом можно судить из его анализа чего-то подобного, робота-зомби с механической памятью. «Такой робот мог прийти к своему нынешнему состоянию... разными путями. Но, независимо от того, как он пришел к нему, он, будучи чисто меха-

нической 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» [Васильев 2009, 221]. Наличное состояние робота может отражать прошлые события. Но, по мнению российского философа, механическая память робота не чувствительна к его фактической индивидуальной истории. На чем основана такая уверенность?

Возьмем убеждение, что я родился в 1976 г., и запись компьютера о том, что первый файл на его диск записан в 1976 г. В чем принципиальная разница? Оба «воспоминания» содержат хронологическую метку, оба могут быть верными или ошибочными, оба в качестве основания имеют некоторый правдоподобный источник (часы реального времени и свидетельство о рождении), и оба влияют на поведение. Оба «воспоминания» могут как-то отражать другие события, и оба могут исчезнуть или измениться, если оказать непосредственное воздействие на физический носитель, мозг или жесткий диск. Повреждения или значимые изменения в физическом носителе в обоих случаях приводят к изменениям или исчезновениям воспоминаний.

Может, различие в том, что механизм компьютерной памяти относительно хорошо изучен, а механизм работы человеческой памяти все еще отчасти остается загадочным для ученых? Но отсутствие понимания всех механизмов работы человеческой памяти не является доводом в пользу нали-

чия у нее нефизических свойств и отсутствия локальной супервентности. Или различие в том, что память человека может содержать конкретные визуальные образы прошлого, а память компьютера – только последовательность битов, которые могут быть проинтерпретированы самым различным образом? Например, запись о времени создания файла какой-то другой программой действительно может быть проинтерпретирована иначе: как математическая формула, данные об изображении, команда распечатать файл и т. п. С этим возражением можно работать (можно попытаться показать, что воспоминания человека могут также быть интерпретированы по-разному, как и любая сжатая информация; тактики атаки и защиты подобных утверждений присутствуют, например, в дискуссии Сёрла и Деннета о «Китайской комнате»). Но более простой способ атаковать АКТ 1 – заменить цифровой жесткий диск в Аргументе-двойнике на аналоговый носитель, например аналоговую кассету или диафильм. Эти носители сохраняют информацию в максимально конкретном виде, и альтернативная интерпретация данных оказывается затруднена. Мне кажется, что параллельный аргумент и его несостоятельность должны показать то, что сам по себе АКТ 1 недостаточен для опровержения локальной супервентности. Для опровержения этого стороннику АКТ 1 нужно показать, что аргументы неэквивалентны.

И в рассуждениях Васильева есть указание на еще одно, возможно, ключевое отличие механической памяти от па-

мента человека. Видимо, именно оно должно делать АКТ 1 правдоподобным, а его механическую вариацию, АКТ 2, – нет. В представлении автора память человека – это приватные качественные состояния. И поэтому «индивидуальное прошлое не растворяется в нейтральных нейронных кодах, а продолжает существовать на уровне ментальных образов» [Васильев 2009, 221]. Это утверждение нужно использовать как еще одну посылку в аргументе. Она фигурирует там проходящим образом, как нечто само собой разумеющееся, но крайне важна и потому требует разъяснения и оценки.

Квалиа – термин, обозначающий качественный или феноменальный аспект ментальных процессов: «мучительность» боли, «красноту» красного, «горечь» обиды, «сладость» наслаждения. Его используют для того, чтобы провести различие между содержательным аспектом ментального процесса (контентом) и субъективным переживанием этого процесса, «каково это: испытывать» процесс на себе. Когда Васильев говорит о том, что человеческая память состоит из приватных качественных состояний, он имеет в виду наличие этого качественного аспекта. Вот я вспоминаю прошлое лето, выходные: дача, на траве стоят плетеный столик и кресло, книги и чашка чая на столе, почти ощущаю запах скошенной травы, слегка греет солнце, чувствую удовольствие с оттенком сожаления. Воспоминания как бы оживляют пережитые ощущения, мои личные, те, что никто другой, кроме ме-

ня, не переживал и о которых другие могут узнать только косвенным образом. Этот пример должен объяснить утверждение Васильева. Но если хорошо подумать, все ли содержимое памяти представлено таким же образом? Вот некто знает иностранные языки, то есть помнит тысячи слов, выражений и грамматических правил. Разве такое содержание памяти окрашено феноменально? Каково это: «помнить три иностранных языка»? И отличается ли это состояние от состояния «помнить четыре или пять иностранных языков»? Сейчас, когда я печатаю этот текст, я, очевидно, помню собственную биографию, но каково это: «помнить свою биографию»? Отличается ли это состояние от состояния «помнить биографию Стива Джобса или Бенджамина Франклина»? А было бы мое состояние другим, если бы вместо того, чтобы сейчас помнить, что я родился в Москве, я, например, помнил, что родился в Бейруте или Бостоне? Память об этих моментах, кажется, влияет на мое поведение. Но чтобы иметь кажимость влияния, они не должны быть как-то качественно окрашены.

Теперь попробуем расставить все по местам. У нас есть АКТ 1 и АКТ 2. Единственным различием между ними является допущение, что память человека состоит из качественных частных состояний. Но из наших примеров очевидно, что по крайней мере часть содержания памяти не имеет «качественного вкуса». Должны ли мы, вопреки интуиции, провести различие в каузальной действенности «кавали-

тативно окрашенных» и «квалитативно пустых» воспомина- ний? Или углубиться в интроспективные исследования, что- бы все-таки найти «вкус» всех без исключения записей па- мяти? Мне кажется, однако, что обе альтернативы беспер- спективны. И вернее считать, что память не имеет «квали- тативных качеств». Тем более что такую точку зрения могут занимать не только элиминативисты (противники существо- вания квалиа), но даже такие квалиафилы, как Д. Чалмерс.

В. Память не является квалитативным состоянием

В своей основной монографии, «Сознающий ум» (1996), Д. Чалмерс, один из самых известных современных аналити- ческих философов, утверждает, что ментальное исчерпыва- ется психологическим и феноменальным (квалитативным). Некоторые ментальные свойства сводятся только к психоло- гическим, некоторые – только к феноменальным, а некото- рые имеют двойственный характер. Боль, к примеру, име- ет как психологические, так и феноменальные аспекты. Пси- хологической составляющей является состояние, связанное обычно с повреждением организма и ведущее к особому по- ведению, реакции уклонения. Как мы видим, это описание функциональное – оно фиксирует входящие параметры и со- ответствующее этим параметрам поведение-реакцию. Фено- менальным аспектом боли является ее мучительное пережи-

вание. Концептуально это не связанные аспекты: теоретически организм мог бы быть поврежден и демонстрировать характерное для боли поведение, но при этом не испытывать болезненных переживаний, и наоборот, мог бы быть не поврежден и не демонстрировать боли в поведении, но интенсивно ее переживать. Но фактически в мире и в жизни организма они переплетены.

А вот память Чалмерс относит целиком к психологическим процессам: «Согласно нашему анализу в параграфе 3, обучение и память – *функциональные свойства*, характеризующиеся каузальными ролями...» (курсив мой. – Д. В.) [Chalmers 1997, 24]. И причина в том, что процесс запоминания и хранения в памяти можно легко представить неосознаваемым, лишенным феноменальных свойств. Почему тогда автор Аргумента каузальных траекторий считает, что память состоит из приватных качественных состояний?

Вероятно, Васильев, когда пишет о памяти, имеет в виду не процесс запоминания и хранения информации о прошлом, а воспроизведение этой информации в текущем опыте субъекта. Только в текущем опыте часть воспоминаний действительно представляется феноменально окрашенной. Но тогда точнее будет сказать, что не память обладает качественностью, а текущее состояние, объектом которого являются события из прошлого. Содержание памяти может быть воспроизведено в сознании с феноменальными свойствами, но не имеет их в процессе хранения. Качественные аспек-

ты проявляются в наличном состоянии субъекта, а не в его памяти.

А что, если представить, будто эти ментальные содержания все-таки живут внутри нас в виде ментальных образов, постоянно текущих феноменальных состояний? Непонятно, как это представить, и непонятно, чьими бы эти состояния были. Если сейчас, когда я пишу эти строки, продолжает существовать феноменальное воспоминание о моем воскресном дне летом, то не совсем понятно, где и для кого оно продолжает существовать. Ведь онтология феноменальных состояний субъективна: если феноменальные состояния существуют, они существуют представленными для субъекта. Уверен, что в нейронных сетях закодированы воспоминания, но существование не осознаваемых никем феноменальных состояний мне представляется маловероятным. Тем более это трудно представить в связи с прерывностью сознательного опыта. Даже если прошлые феноменальные состояния продолжают существовать где-то на периферии моего сознания, то они должны улетучиваться во время и после наркоза или даже сна. Но было бы странно предполагать, что только ментальные события после сна имеют каузальную действенность, а все, что происходило до сна, «растворилось в нейронных кодах». Чтобы защитить аргумент, недостаточно показать возможность «проигрывать» в настоящем воспоминания в «театре сознания». Нужны доводы в пользу непрерывного параллельного существования фено-

менальных состояний прошлого и непротиворечивый способ помыслить это. А этих доводов, судя по всему, мы пока не находим, и неосознанные сознательные состояния помыслить невозможно. Таким образом, различие в квалиа представляется мне нерелевантным. И мы возвращаемся к исходной ситуации: если бы работал АКТ 1, работал бы и АКТ 2.

С. Решение парадокса сближающихся в ментальных состояниях двойников

Впрочем, может быть, и тезис о супервентности ментального на физическом способен привести к трудностям подобного характера? Кажется, такое сомнение озвучивает Васильев, когда предлагает нам представить двойников со сближающимися физическими состояниями. Вот в разных мирах существуют два человека с разными физическими состояниями, но их состояния постепенно сближаются. По мере этого сближения должны постепенно стираться и их ментальные различия. Если это напоминает амнезию, стирание тех аспектов прошлого, которые отличали их друг от друга, двойникам «придется забыть очень многое или даже почти все, если двойников очень много, а ведь это тоже реально возможно» [Васильев 2009, 219]. Чем больше двойников с альтернативными траекториями, тем больше воспоминаний должны быть стерты. Чем больше двойников, тем меньше каждый из них может помнить. И человек, у которого

бесконечное множество двойников, не должен помнить ничего. А значит, «сознанием в 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» [Васильев 2009, 219]. Такое следствие действительно похоже на абсурд. Оно явно не согласуется со здравым смыслом и естественными установками. Неужели мы вынуждены принимать его, если отстаиваем Тезис о супервентности и допускаем множественность каузальных траекторий?

Для того чтобы ответить на этот вопрос, нужно представить ситуацию с двойниками с большими подробностями. Итак, есть два двойника, два почти неотличимых по физическому строению человека, живущие в разных мирах. Пусть один из них, Дэн, родился в Бостоне, а другой, Фил, в Бейруте. И у них разные занятия: Дэн бизнесмен, а Фил – музыкант. Один проводит все время в офисе, другой – играет джаз на фортепиано в маленьком ресторанчике. В какой-то момент параллельных, но не пересекающихся жизней мозги двойников приходят в одинаковые состояния, одинаковые вплоть до каждого нейрона, атома, элементарной частицы. И в этот момент их спрашивают в каждом из миров: «Эй, ты кто? и чем ты здесь занимаешься?» Согласно ТС их ответ должен быть идентичным. Предположим, они оба отвечают: «Чего тебе? Я музыкант, я играю на фортепиано». Но как такое помыслить? Как бизнесмен мог дать такой ответ?

Неужели он забыл все те часы, которые провел в офисе? Он действительно должен был забыть многое, очень многое. Но очень многое он должен был и «помнить». К примеру, он должен помнить мелодию и гармонии джазовых стандартов «Каравана» и «Мисти», неподражаемые виртуозные пассажи легендарного Арта Тэйтума, должен тонко чувствовать свинговый ритм, синкопы, уметь импровизировать. Разве похоже все это на поведение и состояния зомби? Разве его нельзя назвать полноценным сознательным существом? Допустим, кроме Дэна и Фила были и другие двойники, и все они пришли к тому же состоянию, и все теперь говорят, что джаз – их призвание. Разве от этого сознание Дэна или Фила станет менее богатым? Вряд ли нам нужно так думать. Парадокс работает только в том случае, если допустить, что они только забывают, но не приобретают новые воспоминания. Только тогда это вело бы к сокращению возможного содержания памяти.

Действительно ли утверждение о том, что двойники по мере схождения их физических состояний должны забывать отличающие аспекты прошлого, ведет в итоге к сокращению возможного содержания памяти? Мне кажется, это могло бы быть истинным только в случае конечного числа возможных воспоминаний. В самом деле, если от конечного числа возможных воспоминаний отнимать все большее и большее их количество, будет оставаться меньше возможностей. Но если количество бесконечно, то от вычитания все большего и

большого числа различий мощность исходного множества не уменьшится. Хорошо, тогда, может быть, бесконечное число двойников будет представлять опасность для содержания памяти? Мне кажется, что и здесь не возникает проблемы. Ведь если из бесконечного множества натуральных чисел удалить бесконечное множество нечетных чисел, останется бесконечное множество четных. И это множество по мощности будет таким же, как исходное множество.

Но, по моему мнению, абсурдной ситуации (уподобления двойников зомби) не возникнет, даже если количество возможных воспоминаний конечно. Дело в том, что двойники не только «забывают» отличающие их аспекты прошлого, но и приобретают новые воспоминания. Так, двойник-бизнесмен в нашей ситуации не только забыл часы, проведенные им в офисе, но, судя по всему, и приобрел множество дополнительных воспоминаний, связанных с игрой на фортепиано. Содержание его памяти не уменьшилось, а только изменилось. И если физически состояние его мозга сближалось с состоянием музыканта постепенно, стоит предположить, что постепенно менялось содержание его памяти. Постепенно он забывал часы, проведенные в офисе, и убеждался, что вместо офиса проводил время в ресторане и что смотрел он не в экран компьютера, а в ноты. Хотя все это представить очень сложно, зная о сложности устройства мозга, не менее сложно представить вероятность того, чтобы два мозга оказались в абсолютно тождественном состоянии.

Из рассуждений, приведенных выше, следует, что возможно непротиворечиво помыслить двойников со сближающимися ментальными траекториями. И что Тезис о супервентности не ведет к противоречивым последствиям. Это означает, что локальный интеракционизм, скорее всего, не дает бесспорного решения проблемы ментальной каузальности. Но тогда разве не возвращаемся мы к исходной дилемме? После критики локального интеракционизма разве не вынуждены мы отказываться от ментальной каузальности вообще? Или придется соглашаться с тождеством ментального и физического? Обе жертвы крайне нежелательны, а это значит, следует искать другое решение.

IV. Решение проблемы ментальной каузальности в биологическом натурализме Дж. Сёрла

1. Ключевые положения биологического натурализма Дж. Сёрла

Альтернативный выход представляет позиция американского философа Дж. Сёрла. Его теория, биологический натурализм, также предлагает вариант решения проблемы ментальной каузальности и сформулирована им в ряде монографий⁹. Биологический натурализм похож на локальный интеракционизм тем, что в нем также отрицается тождественность ментального и физического. Сёрл тоже настаивает на нередуцируемой природе ментального¹⁰. И также утверждает каузальную эффективность ментальных состояний. Но проблему ментальной каузальности американский философ

⁹ См. [Сёрл 2002; Сёрль 2004; Searle 1997, 1998, 1999, 2004, 2006, 2008 и др.].

¹⁰ В одной из своих монографий Сёрл пишет: «Я настаиваю на том... что мы можем принять очевидные факты физики... не отрицая в то же самое время очевидные факты нашего внутреннего опыта – например, что все мы обладаем сознанием и что наши сознательные состояния имеют достаточно специфические *нередуцируемые* феноменологические свойства» (курсив – авторский) [Searle 1994, 28].

пытается решить по-другому, не ставя под сомнение супер-превентность ментального на физическом. В этом – ключевое отличие его позиции от локального интеракционизма. Позицию Дж. Сёрла в отношении сознания и ментальной каузальности можно резюмировать в следующих тезисах:

Онтологический статус сознания: Ментальные состояния существуют как высокоуровневые характеристики, но нередуцируемы к низкоуровневым.

Ментальная каузальность: Ментальные состояния каузально влияют на наше поведение.

Связь с состояниями мозга: Ментальные состояния полностью каузально обусловлены процессами в мозге.

Каузальная замкнутость: Принцип каузальной замкнутости работает, но Принцип каузальной исключительности не работает.

Первые два тезиса, как мы видели, являются общими для локального интеракционизма и биологического натурализма. Нередуцируемость ментального для Сёрла обусловлена исключительными характеристиками ментальных феноменов. Всего Сёрл перечисляет четыре особенных свойства ментальных состояний: (а) субъективность, (б) качественный характер, (в) интенциональность и (г) единство опыта [Searle 1994]. Под субъективностью ментального имеется в виду именно онтологическая, не эпистемическая субъективность. Сёрл пишет: «ментальные состояния обладают нередуцируемой субъективной онтологией» [Searle 1994, 19].

Для Сёрла ментальные феномены существуют только как часть внутреннего опыта человека (или животного).

Когда Сёрл говорит о качественном характере, он имеет в виду качественный характер внутренних состояний. По его мнению, ментальное необходимо имеет этот уникальный для каждого субъекта характер, «то, что значит быть» в конкретный момент этим самым субъектом. К примеру, ощущение вкуса пива, запаха свежеспеченного хлеба, звуков Третьей симфонии Бетховена – все это внутренние, непосредственно данные субъекту и недоступные для других переживания. Под интенциональностью Сёрл подразумевает направленность состояний на объекты: радость чему-то, сожаление о чем-то, ожидание чего-то, надежда на что-то. Почти все ментальные состояния имеют такое свойство. И наконец, единство опыта – это объединение в опыте зрительных образов, тактильных ощущений, звуков, запахов и проч. Сёрл подчеркивает здесь холистическое свойство сознательного опыта¹¹.

Итак, каузальная эффективность ментального тоже утверждается обоими философами. Только российский философ считает это утверждение выводом из истинных посылок, а Сёрл – очевидностью. По мнению Сёрла, ментальная каузальность – это эмпирический факт, с которого и нужно начинать. «Я думаю, что мы ощущаем необходимую [каузаль-

¹¹ Резюмирующее описание теории Сёрла можно найти в его статье «Биологический натурализм» [Searle 2007].

ную] связь почти всю нашу сознательную жизнь...» [Searle 2004, 203]. Вот, говорит он на очередной лекции, я хочу поднять руку, и проклятая штука поднимается, поднимается из раза в раз. Только чудак способен сомневаться в том, что причиной движения руки является желание.

Сёрл удивительным образом сочетает тезисы о ментальной каузальности и каузальной замкнутости физического мира. В третьем из выделенных выше тезисов он утверждает, что ментальные состояния с их специфическими характеристиками вызваны процессами в мозге и только ими. Но это еще ничего не говорит о том, чем именно являются ментальные состояния. Может даже показаться, что Сёрл считает, будто процессы в мозге являются причиной каких-то других процессов, происходящих не в мозге, а где-то еще, в духовной субстанции например. Но это не так: его позиция по этому вопросу зафиксирована в первом тезисе. Ментальные состояния реализованы в самом мозге как высокоуровневые характеристики системы.

В качестве аналогии высокоуровневой характеристики Сёрл приводит такое свойство, как текучесть. Одна молекула H_2O не обладает свойством текучести, но множество молекул – обладают. При этом текучесть не является мистическим или необъяснимым феноменом. Она складывается из свойств скоплений молекул. Причиной текучести, согласно Сёрлу, являются свойства скопления молекул, при этом текучесть «реализована» в скоплениях молекул. Аналогичным

образом, по его мнению, обстоит дело с сознанием. Причиной ментальных состояний являются процессы в мозге, при этом сознание реализовано в мозге.

Минимальным расхождением между локальным интеракционизмом и биологическим натурализмом представляются основания второго тезиса. Васильев считает, что ментальная каузальность требует доказательства, а Сёрл, кажется, видит возможность обосновать ее простой демонстрацией. Является ли ментальная каузальность эмпирическим фактом или гипотезой, требующей доказательств?

Думаю, позиция сторонника локального интеракционизма здесь более верна. Действие желания как причины мы не можем наблюдать. Ну и что, что после того, как мы ощущаем желание поднять руку, «чертова штука» поднимается из раза в раз? Ментальная каузальность – не факт, не очевидность, а гипотеза, которую нужно доказывать. Такой вывод следует из открытий Д. Юма. Фактом является не наличие причинно-следственной связи, в частности, между желанием и действием, а регулярное следование одного за другим во времени. Вот я хочу поднять руку, и вот – она поднимается. Если бы фактом было наличие причинно-следственной связи между желанием и действием, фактом была бы и причинно-следственная связь между днем и ночью. Вот день, вот ночь. Они всегда следуют друг за другом. Но разве день является причиной ночи? Причиной наступления дня и ночи является нечто иное – движение планеты вокруг своей

оси. Даже если последовательность есть, причина может находиться в другом месте. Так может обстоять дело и с ментальной каузальностью. Последовательность является эмпирической данностью, но необходимую связь, каузальную эффективность нужно доказывать.

Теперь разберемся в пунктах, в которых сосредоточены основные различия. Сёрл утверждает, что ментальные состояния полностью обусловлены процессами в мозге. То есть что ментальные состояния детерминированы состояниями мозга и только ими. Но какого рода эта связь? Сёрл считает, что ментальные состояния детерминированы мозгом, потому что они *в качестве причины* имеют физические процессы в мозге и реализованы в мозге в качестве высокоуровневой характеристики. «Обобщая... мы могли бы сказать, что два феномена могут быть как связаны каузальной связью, так и реализованы один другим, если они находятся на разных уровнях описания» [Searle 1999, 266]. Это высказывание оригинально и даже удивительно потому, что сочетает в себе две, казалось бы, противоположные идеи. Сёрл как бы смешивает здесь понятие причины и субстрата. Несколько необычно утверждение, что нечто может быть реализовано в чем-то и это что-то одновременно может быть его причиной. Но не только соотношение последних двух тезисов должно вызывать удивление. Как согласуются утверждения, что сознание есть высокоуровневая характеристика мозга и что при этом оно не редуцируется к состояниям мозга? Ведь

высокоуровневые характеристики должны редуцироваться к характеристикам низкого уровня, а Сёрл утверждает нередуцируемость ментального.

Концепция Сёрла нарушает различие, которое мы провели в начале исследования: между расшифровывающей функцией подчинительного союза «потому» и каузальной. Кажется, Сёрл демонстрирует, что союз «потому» может указывать сразу на оба отношения: и концептуальное, и каузальное. Примеры самого Сёрла лучше прояснят его позицию.

2. Иллюстрации к каузальным связям в биологическом натурализме

Рассмотрим переход воды в газообразное состояние в процессе кипения и проанализируем его с позиции Сёрла¹². Если отбросить связь с агентом (пока оставим ментальную каузальность в стороне и рассмотрим только очевидно физические процессы), кипение можно объяснять так: вода закипает потому, что разогревается зажженной конфоркой до соответствующей температуры. То есть (1) разогрев воды на огне приводит к переходу ее в газообразное состояние. Кажется, это вполне законное каузальное объяснение. Но без труда найдется и другое описание, которое претендует на роль каузального: кипение происходит потому, что кинети-

¹² Этот пример Сёрл рассматривает в книге «Открывая сознание заново» [Searle 1994, 87].

ческая энергия, передающаяся в результате окисления углеводородов (горения топлива) молекулам H_2O , заставляет их двигаться так быстро, что между ними разрываются (полярные и водородные) связи. То есть (2) повышение кинетической энергии до определенного уровня приводит к разрыву связей между молекулами воды. Таким образом, у нас два совершенно законных каузальных объяснения перехода воды в газообразное состояние. Первое – на макроуровне (макрособытие $A \rightarrow$ макрособытие B), а второе – на микроуровне (микрособытие $A \rightarrow$ микрособытие B). Оба описания позволяют проследить связь между антецедентом и консеквентом. Эти описания отражают последовательность событий во времени, поэтому вслед за Сёрлом назовем их описаниями «слева направо» [Searle 1994, 87]. Есть ли между этими описаниями конкуренция? По мнению философа, в данном случае проблему конкуренции, даже если она возникает, решить не сложно. Достаточно провести соотношение между событиями микро- и макроуровней. То есть идентифицировать разогрев воды с повышением уровня кинетической энергии молекул воды, а газообразное состояние воды – с состоянием, когда молекулы воды не объединены полярными и водородными связями между собой. Как только идентификация осуществлена, проблема исчезает: у нас не две различные причины, не два конкурирующих за каузальную эффективность явления, а одно, но с двумя описаниями на двух разных уровнях.

Наверное, тогда можно добавить и другие каузальные описания, тоже без конкуренции: микро-макро и макро-микро слева направо: (3) повышение кинетической энергии молекул воды до определенного уровня является причиной кипения, (4) разогрев воды до определенного уровня приводит к разрушению полярных и водородных связей. Таким образом, есть уже четыре способа каузального описания слева направо: микро-микро, микро-макро, макро-макро и макро-микро [Searle 1994, 87]. Эти способы возникают в силу того, что одно и то же явление можно описывать на разных уровнях. И, по всей видимости, конкуренция за каузальную эффективность между этими описаниями снимается через идентификацию явлений нижнего и верхнего уровней.

Но с позиции Сёрла в описанном явлении существует еще один тип каузальных связей – «снизу вверх». Этот тип связи наиболее экстравагантен. Судя по его рассуждениям¹³, Сёрл утверждает наличие каузальности между одномоментными, синхронными событиями на разных уровнях. То есть повышение кинетической энергии молекул воды он также считает причиной повышения температуры воды. И соответственно, (5) разрыв связей между молекулами воды является причиной перехода воды в газообразное состояние. Таким образом, по мнению Сёрла, правомерно еще и пятое описание

¹³ Эти рассуждения Сёрла о ментальной каузальности приведены в Эпилоче к его монографии «Интенциональность. Эссе по философии сознания». Вместо нашего примера с кипением воды Сёрл в этом тексте анализирует работу двигателя внутреннего сгорания, тоже на двух уровнях [Searle 1999, 262–272].

причины кипения воды: снизу вверх, микро-макро. Конкуренция в этом случае, по его мнению, также не является проблемой. Ведь речь идет об одном и том же явлении.

Этот пример является важным инструментом для иллюстрации подхода к феномену сознания и решения проблемы ментальной каузальности Сёрлом. Он, наверное, должен служить доводом в пользу существования множественности достаточных причин для простых физических процессов. Теперь остается провести следующую аналогию между соотношением высокоуровневых и низкоуровневых характеристик физических процессов – с одной стороны, и соотношением ментальных и нейронных процессов – с другой. Сёрл ее и предлагает.

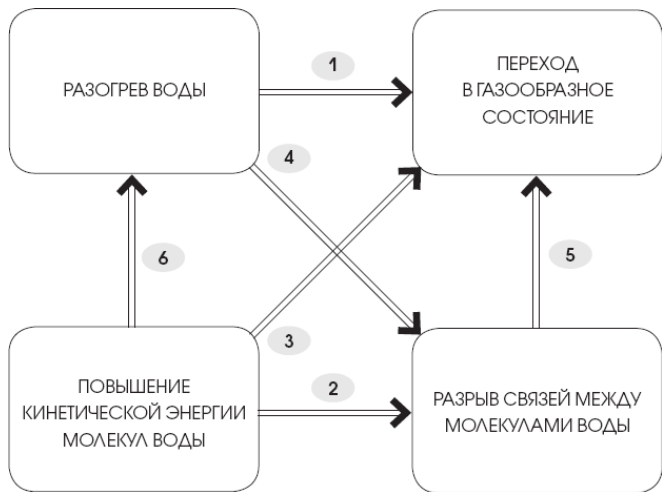


Рис. 1. Схема каузальных связей процесса кипения воды по Сёрлу

Философ утверждает, что отношение между мозгом и сознанием, физическими и ментальными состояниями, является не чем-то удивительным, а вполне тривиальным. И что мозг является причиной ментальных состояний, так же как причиной этих ментальных состояний являются предшествующие ментальные состояния. Мозг каузально влияет на сознание, и сознание – на последующие состояния сознания. Таким образом, Сёрл вписывает ментальное в каузальные цепочки природы и представляет конфликт ментальной и физической причин иллюзией.

Рука поднимается вверх. Причиной этого действия, по его мнению, являются как желание поднять руку агентом, так и возбуждение такой-то сети нейронов в его мозге. Но это – не две конкурирующие причины. Желание поднять руку и возбуждение сети нейронов – это описания события на разных уровнях: желание – на макроуровне, а возбуждение нейронов – на микроуровне, подобно тому, как разогрев воды является описанием процесса на макроуровне, а повышение кинетической энергии молекул воды – на микроуровне.

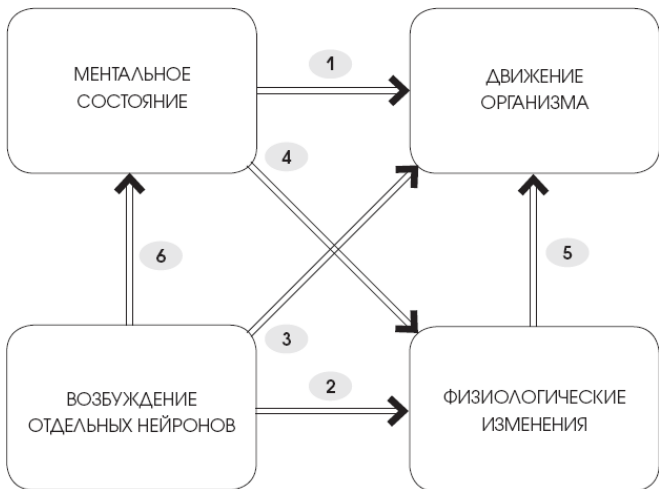


Рис. 2. Схема каузальных связей движения организма по Сёрлу¹⁴

Рассуждения Сёрла вполне понятны. Полезным кажется его указание на возможность параллельно описывать явления на макро- и микроуровнях. Но (1) действительно ли когерентно такое описание причин физических явлений? И (2) действительно ли проходит аналогия между высокоуровневыми и низкоуровневыми характеристиками физических процессов, с одной стороны, и нейрофизиологическими процессами и состояниями сознания – с другой, в рамках теории сознания Сёрла?

¹⁴ Подобный рисунок Сёрл приводит в том же эпилоге [Searle 1999, 270].

3. Внутренние противоречия теории Дж. Сёрла

Само описание причинных связей в примере Сёрла вызывает у меня сомнение, в особенности 5-е направление каузального действия – синхронного. Причиной газообразного состояния воды, по мнению Сёрла, является свободное движение молекул. Но ведь газообразное состояние и есть состояние, в котором молекулы движутся свободно. Причиной жидкого состояния воды философ называет ограниченное движение молекул в условиях полярных и водородных связей. Но именно ограниченное движение молекул в условиях полярных и водородных связей и является жидким состоянием. Между двумя парами этих состояний имеет место редуکتивно-супервентное отношение явлений разного уровня или логическая связь, а не каузальная. То есть газообразное состояние воды редуцируется к свободному и быстрому движению ее молекул и является супервентным по отношению к свободному и быстрому движению молекул. Проблема для Сёрла возникает в связи с симметрией описываемых отношений. Философ представляет здесь каузальные отношения как симметричные, но в таком случае понятие причинно-следственной цепи становится, на мой взгляд, некогерентным¹⁵. Причинная связь асимметрична, причины опре-

¹⁵ К подобным выводам приходит в своей диссертации и А. В. Кузнецов. Он

деляют действия, а не наоборот.

Предположим, что в момент t_1 происходят низкоуровневые события. И в этот же момент происходят высокоуровневые события. Изменения высокоуровневых событий не происходят без изменений низкоуровневых событий. Существует очевидная регулярность. Если допустить, что низкоуровневые события являются причинами, а высокоуровневые события – следствиями, то по той же самой логике высокоуровневые события являются причинами, а низкоуровневые события – следствиями. Следствие является причиной, а причина – следствием. Порочный круг можно разорвать, но только допустив асимметрию в отношении разных уровней свойств. Скажем, утверждая, что каузально эффективными всегда являются низкоуровневые свойства. В пользу этого хода можно привести доводы. К примеру, оправданно свойства целого выводить из свойств его частей. Физика считается самой фундаментальной наукой именно потому, что изучает элементарные частицы и законы их взаимодействия. Но именно для Сёрла этот ход закрыт. Ведь сам философ неоднократно утверждает возможность каузального действия не только «снизу вверх», но и «сверху вниз». См., например, рис. 1, связь 4. Понятие каузальной зависимости, мне кажется, предполагает асимметричность, а Сёрл ее нарушает. По-

пишет: «Понятие эмерджентной причинности [Сёрла] противоречит определению причины и действия, которые должны быть разведены по времени. События, происходящие одновременно, не могут быть по отношению друг к другу причиной и действием согласно дефиниции» [Кузнецов 2016b, 109].

этому приведенный способ каузально описывать процесс кипения мне кажется некогерентным.

4. Недостатки проведенных аналогий и иллюстраций

Думаю, что и аналогии Сёрла работают не совсем удачно – с ними возникает проблема именно в рамках биологического натурализма. Аналогия между ментальным и высокоуровневыми характеристиками физических процессов плохо согласуется с другими положениями биологического натурализма, да и не схватывает специфических свойств ментального по теории Сёрла (субъективность, качественность, интенциональность, единство). Высокоуровневые характеристики физических процессов (например, кипение) действительно реализованы в низкоуровневых микроструктурах. Но в данном случае они редуцируемы к характеристикам микроструктур, совпадают в пространстве и выводимы из движения микроструктур. Наблюдая, скажем, поведение молекул воды, можно было бы предположить свойства газообразного, жидкого и твердого состояний. И переход от микро- к макроуровням можно проследить эмпирически.

Проведем исследование растекающейся жидкости под микроскопом. Сначала на максимальном увеличении наблюдаем движение одной молекулы воды. Если мы помениаем кратность увеличения, то увидим несколько молекул,

их скопление и общее движение. Если еще уменьшить кратность, то, наконец, станет видно движущуюся каплю воды, то есть каплю, обладающую свойством текучести. Хотя это – увлекательное исследование, оно не шокирует. Изменяя степень увеличения, мы тем не менее не изменяли перспективу. Обнаруживая, как скопления движущихся молекул формируют каплю, мы не обнаруживаем в какой-то момент, что не наблюдатель рассматривает каплю, а капля – радужку и зрачок глаза. Однако именно такой радикальный скачок получился бы, если бы, по логике Сёрла, изменяя параметры оптики наблюдения нейрофизиологических процессов, мы вдруг увидели бы онтологически субъективные ментальные состояния агента.

Мне кажется, что аналогия работала бы только в случае, если бы Сёрл отождествлял ментальные процессы с физическими. Ведь высокоуровневые характеристики разогрева и кипения воды можно отождествить с низкоуровневыми процессами увеличения кинетической энергии молекул воды. Но в теории сознания Сёрла нельзя ожидать такого отождествления. Ведь среди его ключевых тезисов – утверждение о нередуцируемости ментального и его уникальных специфических черт, в первую очередь – субъективности. Сознание, согласно его мнению, онтологически субъективно, то есть существует как приватные субъективные состояния. Но это принципиально отличает сознание от физического. И следовательно, аналогия в теории Сёрла не проходит. Из все-

го этого можно сделать вывод о том, что решение проблемы ментальной каузальности в биологическом натурализме не оптимально. Альтернативой является теория коллеги Сёрла и его наиболее знаменитого оппонента, философа Д. Деннета.

V. Решение проблемы ментальной каузальности в телеофункционализме Д. Деннета

1. Ключевые положения телеофункционализма Деннета

Д. Деннет является одним из самых важных и известных современных аналитических философов. Он – не только автор множества монографий и сотен статей, но и сам герой множества книг. Главный вклад Деннета в философию сознания заключается в том, что он заставил коллег отвлечься от исключительно «кабинетных изысканий», показал, что исследования в области когнитивистики, нейрофизиологии и биологии имеют отношение к обсуждению философских проблем¹⁶. Разработанный Деннетом подход можно назвать телеофункционалистским (ТФ). «Я, конечно же, – своего рода “телеофункционалист”, возможно, даже первый телеофункционалист» [Dennett 1991b, 460], – пишет он в монографии «Объясненное сознание». Теория сознания Деннета оснащена множеством эмпирических положений, но в ней

¹⁶ Этот призыв Деннет сделал в статье «От кабинетной философии – в поле» [Dennett 1998, 289].

можно выделить философскую основу. Метафизические положения телеофункционализма сводятся к следующему:

Онтологический статус сознания: Ментальные состояния существуют как высокоуровневые характеристики физических свойств.

Ментальная каузальность: Ментальные состояния каузально влияют на наше поведение.

Связь с состояниями мозга: Ментальные состояния полностью реализованы процессами в мозге.

Каузальная замкнутость: Принцип каузальной замкнутости работает в физическом мире, но некоторые события могут быть объяснены только высокоуровневыми закономерностями.

Деннет согласен с Сёрлом в том, что ментальные состояния представляют собой высокоуровневые характеристики физических свойств. Однако в отличие от Сёрла он не считает, что эти свойства – биологического характера. Деннет убежден, что любой механизм, обладающий функциональными качествами мозга, будет обладать ментальными состояниями [Dennett 1991b, 281]. Как и Сёрл, Деннет уверен, что ментальные высокоуровневые характеристики каузально эффективны. Однако в отличие от Сёрла он не считает, что мозг является причиной появления сознания. По его мнению, ментальные свойства реализованы в физическом субстрате как высокоуровневые – в низкоуровневых. Деннет также, судя по всему, разделяет Принцип каузальной замкнутости физического мира, но допускает наличие высо-

коуровневых причин. Скорее всего, он не разделяет Принцип исключения, хотя в открытом виде этого утверждения я у него не нашел.

Телеофункционализм Деннета является современным развитием функционалистской программы объяснения сознания. Функционализм – это позиция, согласно которой сущность ментальных состояний выражается в способе их функционирования, в той роли, которую они играют в других процессах. Определить наличие ментального содержания можно по результатам обработки системой входящих сигналов. Например, боль с функционалистской позиции – это состояние, ведущее к реакции избегания; симпатия – наоборот, к притяжению. Убеждение в том, что «сегодня будет дождь», – это что-то вроде предрасположенности взять зонтик или плащ после просмотра прогноза погоды в случае пребывания в состоянии «желания остаться сухим». Если обобщить, то по этой логике *убеждения* представляют собой состояния, определенные данными органов восприятия, а также другими убеждениями, и при взаимодействии с желаниями формирующие поведение или действия. А *желания* – это состояния, находящиеся в связи с другими желаниями и намерениями, при взаимодействии с убеждениями формирующие поведение или действия.

Функционализм предполагает возможность множественной реализации¹⁷ ментальных состояний. То есть, по мнению

¹⁷ Это понятие ввел в философию сознания Х. Патнэм. В серии статей, направ-

функционалистов, субстрат и внутренняя механика системы не определяют ментальных содержаний. Биологический носитель – только один из вариантов. Вполне возможно ментальные содержания реализовать и на механических носителях, в частности на кремниевых чипах. Такой подход характеризует все виды функционалистских теорий. Взгляды Деннета тем не менее имеют некоторую специфику.

Позицию Деннета можно для контраста сравнить с другим видом функционализма – с функционализмом машины Тьюринга (ФМТ)¹⁸. Эта теория несколько более ранняя и, возможно, менее совершенная. Для иллюстрации общности и различий используем два ключевых вопроса: (1) что представляют собой ментальные содержания (content)? (2) что общего имеется между двумя системами, когда они имеют одно общее убеждение? Посмотрим, как на них ответят сторонники ФМТ и ТФ.

Ответ на первый вопрос, скорее всего, будет одинаковым для обоих вариантов функционализма. Он служит идентификатором *монизма*. «Каждое ментальное состояние – это какое-то физическое состояние», – считает большинство

ленных против теории тождества, он показал, что ментальные состояния могут быть реализованы различными физическими состояниями, поэтому ментальное и физическое не могут быть в строгом смысле тождественны друг другу [Putnam 1967].

¹⁸ В «Объясненном сознании» Д. Деннет пишет: «Тогда являюсь ли я функционалистом? И да, и нет. Я не функционалист машины Тьюринга» [Dennett 1991b, 460].

функционалистов. А вот по поводу второго вопроса позиция может быть различной. Сторонник ФМТ скажет, что системы, имеющие общее убеждение, могут быть описаны как выполняющие один и тот же алгоритм и находящиеся в одном и том же месте выполнения этого алгоритма. Например, как если бы синхронно запущенная презентация PowerPoint на компьютере IBM и на компьютере Macintosh была остановлена на одном и том же слайде. Сторонник ТФ ответит более аккуратно. Он скажет, что общность между двумя такими системами в том, что с учетом этого убеждения поведение обеих систем (в той мере, в которой оно определяется этим убеждением) можно будет надежно предсказать.

Таким образом, (1) телеофункционалист не требует ни общности в физической реализации, ни общности «программы» (мыслящие существа могут в силу образования, воспитания и наследственных данных значительно различаться «алгоритмами»). Могут быть запущены разные презентации на разных машинах, только бы эти презентации вели к одним и тем же практическим выводам и действиям. В этом смысле телеофункционализм Деннета более либерален, чем ФМТ. Но здесь нужно сразу сделать оговорку. Предъявляя меньшие требования для идентификации ментальных содержаний в отношении алгоритмов, автор теории вводит дополнительное требование: надежное и осмысленное соединение системы с внешним миром.

Деннет не случайно называет свою позицию телеофунк-

ционалистской. Вероятнее всего, это указывает на (2) связь понятий функции и направленности в его теории ментальных содержаний. Для наличия соответствующих интенциональных состояний недостаточно иметь аналогичную реакцию на входящую информацию. Реакция должна соотноситься с функцией системы в целом, ее интересами, приоритетами, целями. В случае с живыми организмами это – биологические, адаптационные функции. Реакция должна быть не просто соответствующей, она должна иметь соответствующий смысл.

Недостаточно, чтобы программа приводила к похожим действиям, она должна осмысленным образом состыковывать интенциональную систему с внешним миром. По мнению Деннета, семантическое содержание приобретает только в целенаправленном взаимодействии с внешним миром. Логика выживания глобально определяет интенциональные содержания человека. Робот тоже может являться системой для выживания и, соответственно, с позиции американского философа, иметь интенциональные содержания. Но у него тогда должны быть цели, интересы, приоритеты. Не все механизмы так состыкованы с внешним миром. Лишенный собственного интереса термостат только очень условно может соответствовать критериям ментальных содержаний. Для человека сохранение температуры тела является принципиальным условием выживания, а для термостата – произвольной функцией. К тому же устройство термостата

не позволяет ему самому определять ответные «действия». Термостат можно использовать как для регулирования потока хладагента, поступающего в радиатор, так и для запуска ядерной ракеты. Биологические сенсорные системы и механизмы управления гораздо избирательней в репертуаре действий, они почти идеально настроены под соответствующие воздействия внешнего мира и обеспечивают адекватную реакцию на стимул.

Но если у мыслящих агентов разные программы, как возможны надежные предсказания? Почему одни мыслящие существа способны «понимать» состояния и предсказывать поведение других? Ключом к разрешению этого вопроса является некоторый стандарт, общая норма, определяющая поведение всех существ. Это стандарт рациональности. И существованием этого стандарта мы обязаны общему эволюционному процессу. Несмотря на то что каждый вид и даже каждая особь могут иметь свою собственную программу рациональности, они все написаны по одному лекалу. У всех программ один общий программист – естественный отбор.

Одна и та же рациональность стоит за строением птичьего крыла и... проектом атомной подводной лодки, созданной человеком. Рациональность, которая повсюду встречается в природе и которая в течение столетий объяснялась креационистской моделью как свидетельство рациональности создателя, объясняется теорией естественного отбора. Из ее основного принципа следует, что если имели место борьба

видов и естественный отбор, то популяции, которые живут сегодня, являются наиболее приспособленными к условиям существования, или обладают *оптимальной конструкцией*. Они выработали наиболее эффективные модели реакций на воздействия среды и способны к стабильному воспроизводству. Общей целью для всех живых существ является сохранение вида, а следовательно, сохранение жизни особи, когда она может произвести потомство. И поведение, которое ведет к достижению этой цели, – наиболее оптимальное или рациональное. В контексте теории эволюции можно еще раз уточнить определения «убеждение» и «желание». Деннет пишет, что «убеждения» и «желания» в конечном итоге демонстрируют необходимость выживания и отражают оптимальное устройство органической системы. На этом и основаны механизмы предсказания поведения агентов. «Если эволюция проделала свою работу... наши предсказания должны быть достаточно надежными, чтобы ими пользоваться» [Dennett 1981, 9].

Любопытно также отметить, что то, что отличает теорию контента Деннета от некоторых функционалистских теорий контента, сближает его позицию с концепцией *интенциональности* – такой, какой она появилась в философском обращении в конце XIX в. Интенциональность (от латинского *intendo* – указывать на что-то, целить во что-то), эта отличительная черта психического, как было подмечено австрийским философом Ф. Brentano, указывает на способ-

ность ментальных актов, с одной стороны, содержать нечто (обладать контентом), а с другой – быть направленными на что-то. Убеждение, к примеру, в том, что Земля круглая, с одной стороны, *содержит в себе* некоторую информацию и отношение к ней, но также, с другой стороны, указывает на некоторый объект за пределами ментального. Как пишет Brentano в книге «Психология с эмпирической точки зрения»: «Каждый психический феномен характеризуется посредством того, что схоласты Средневековья называли интенциональным (а также, пожалуй, ментальным) *существованием в нем некоторого предмета...* и что мы, хотя и не вполне избегая двусмысленности выражений, называли бы отношением к содержанию, *направленностью на объект* (под которым не следует понимать некоторую реальность), или имманентной предметностью. Каждый [психический феномен] содержит в себе объект, хотя и не каждый одинаковым образом. В представлении нечто представляется, в суждении нечто признается или отвергается, в любви – любит, в ненависти – ненавидит, в желании – желает и т. д.» (курсив мой. – Д. В.) [Brentano 1924, 124–125; цит. по: Молчанов 2007, 412–413; см. также: Brentano 1996, 33]. Деннет сохраняет эту двусмысленность термина: в его концепции интенциональность указывает и на содержание, и на направленность. Это – и функциональное репрезентативное состояние, и направленность в мир, на объект.

Деннет, как и Васильев, как и Сёрл, предлагает свое реше-

ние проблемы ментальной каузальности. Однако в данном случае решение Деннета не представляет собой строгий аргумент и уж тем более развернутую теорию. Когда философа спрашивают о ментальной каузальности, он всегда приводит свой мысленный эксперимент, который называется «История двух черных ящиков» [Dennett 1995, 412]. Деннет считает, что описание этой гипотетической ситуации само по себе проливает свет на проблему ментальной каузальности и демонстрирует ее решение.

2. «История двух черных ящиков» как иллюстрация каузальной эффективности высокоуровневых свойств

Представим, что некогда были найдены два соединенных между собой кабелем черных ящика. На первом расположени две кнопки: α и β , а на втором – три световых индикатора: красный, зеленый и оранжевый. Когда стали исследовать ящики, обнаружили, что при нажатии кнопки α на первом ящике на втором всегда загорается красный индикатор, а при нажатии кнопки β – зеленый. Оранжевый индикатор, казалось, вообще не используется. После многократных повторений ученые убедились:

α является причиной красного;

β является причиной зеленого.

Сигнал передавался через провод. Было очевидно, что, когда нажимали кнопку α , в проводе происходило что-то одно, а когда кнопку β – что-то другое. Возможно, различалось количество импульсов тока, а может, разным было напряжение. Для того чтобы выяснить, в чем состояло отличие, ученые подключили к кабелю «прослушивающее» устройство. И вскоре обнаружили, что, когда нажималась любая из кнопок, по кабелю передавался поток импульсов с интервалами, как бы нулей и единиц, всего 10 000 бит. Но каждый раз последовательность была разной. Что в таком случае определяло, какого «цвета» этот поток?

Чтобы разгадать головоломку, ученые вскрыли ящик с цветовыми индикаторами. Там они обнаружили компьютер с огромным процессором, памятью и программой, занимавшей почти весь жесткий диск. Когда на процессор поступала последовательность из 10 000 бит от другого ящика, процессор осуществлял несколько миллиардов операций, в результате которых он выдавал 0 или 1. Если на процессор приходила та же самая комбинация, он выдавал тот же самый результат, 0 или 1, зеленый или красный. Ничего сверхъестественного не происходило. Кроме одного любопытного факта. Компьютер во втором ящике перед выдачей одного и того же ответа на один и тот же запрос проходил через разные физические состояния. Это объяснялось тем, что программа записывала каждый входящий запрос, и состояние памя-

ти компьютера никогда не оставалось прежним. Наблюдая за работой компьютера, ученые решили, что любая последовательность из 10 000 бит будет определяться им как красная или зеленая. Гипотеза, однако, провалилась, как только они стали подменять последовательности битов, исходящие из первого ящика, новыми, собственными. Вдруг заработал оранжевый индикатор, как если бы второй ящик обнаружил мошенничество.

На любую новую последовательность, введенную учеными самостоятельно, загорался оранжевый. Нет, нельзя было сказать, что компьютер как-то отличал последовательности, «сделанные вручную», от последовательностей, исходящих из первого ящика. Когда вручную дублировались последовательности из первого ящика, второй ящик работал нормально, то есть выдавал зеленый или красный. Вот только если изменялся хотя бы один бит, загорался оранжевый индикатор. Выходило, что все возможные последовательности делятся на зеленые, красные и... оранжевые. Только оранжевых было значительно больше, на порядки больше. Это делало закономерность с красным и зеленым индикаторами еще более интригующей. Что все-таки определяло, каким будет «цвет» последовательности символов? В каждом конкретном случае ничего удивительного не было. Ученые могли проследить каузальную цепочку, начиная от ввода любой последовательности в компьютер до вывода им результата. В этом смысле аппарат был полностью детерминирован фи-

зическими законами. Что было удивительно, так это *отсутствие возможности предсказать*, исследуя одну лишь последовательность символов, какой результат выдаст компьютер. Ученые знали лишь, что строка будет оранжевой с почти стопроцентной вероятностью, кроме случаев, когда она была выдана первым ящиком. В последнем случае с вероятностью миллиард к одному она была либо красной, либо зеленой. Но никто *не мог заранее определить один из двух цветов*, не пропустив эту строку через программу в ящике *Б*. Ученые решили, что разгадку даст только вскрытие первого ящика.

Они открыли этот ящик и обнаружили в нем еще один компьютер с гигантской памятью и программой. При нажатии кнопки программа обращалась к встроенным часам, переводила время в бинарное выражение и с помощью этого кода выбирала, к какой части памяти обратиться, чтобы подготовить последовательность для второго ящика. Стало понятно, что именно использование внутренних часов гарантировало, что по кабелю всегда передавалась новая комбинация. При этой псевдопроизвольности всегда, когда нажималась кнопка α , загоралась красная лампочка, а когда β – зеленая. Впрочем, ученые наконец обнаружили аномалию: примерно раз в миллиард случаев происходила «ошибка», и после нажатия кнопки α загорался зеленый цвет. Эта незначительная погрешность сделала аппарат еще более загадочным для ученых.

Но однажды в лабораторию зашли два программиста, Эл и

Бо, которые создали этот аппарат, и все объяснили. Когда-то они оба работали над экспертными системами: создавали базу знаний, в которой должны были содержаться все истинные утверждения о мире и выводимые из них утверждения. Они писали эти программы на разных языках программирования. Оба преуспели в этом деле так, что скоро их программы могли генерировать бесконечное множество истинных утверждений: «Дом больше семечка», «Корова – самка домашнего быка», «Снег бел». Но, как выяснилось, экспертные системы не могли дать полезного нового знания человечеству. И программисты забросили свои проекты, превратив разработки в философскую игрушку.

Они выбрали способ транслировать (переводить) утверждения между машинами и соединили программы так, что при нажатии кнопки α первая машина отправляла истинные утверждения, а при нажатии β – ложные. Вторая машина сопоставляла утверждение на входе с собственными «убеждениями» и, если получала истинное утверждение, подтверждала его красным сигналом, а если ложное – откликалась зеленым. Когда данные на входе невозможно было обработать, то есть когда они выражали бессмыслицу либо содержали некорректные выражения, машина выдавала оранжевый цвет. Конечно, когда ученые вносили коррективы в передаваемые сигналы без понимания того, что они собой должны представлять, это было самой распространенной реакцией. Так раскрылось, что удивительное свойство красных после-

довательностей было свойством «быть истинным утверждением», а зеленых – «быть ложным утверждением». Неожиданно, поняв устройство аппарата, ученые сами смогли создавать сколько угодно зеленых и красных последовательностей. Они даже открыли последовательности, которые сначала были красными, а потом становились зелеными, как, например, сообщение: «это сообщение передавалось менее 100 раз». Ведь на 101-й раз это сообщение должно было становиться зеленым.

3. Интерпретация мысленного эксперимента

Мысленный эксперимент «Два черных ящика» хорош своим почти детективным сюжетом, и вряд ли есть читатель, который хотя бы чуть-чуть не был заинтригован любопытным аппаратом и его свойствами. Однако нас интересует польза, которую философские исследования ментального могут получить от подобной иллюстрации. Не является ли эта громоздкая мысленная конструкция чем-то вроде «машины Руба Голдберга», чрезвычайно сложным и запутанным устройством для осуществления простейшей функции? Быть может, задачу, которую решает мысленный эксперимент, можно решить более лаконичным образом? Да и в чем состоит эта задача?

Эта история, насколько мне известно, была опубликова-

на в двух книгах Деннета: «Опасной идее Дарвина» (1990) и монографии «Помпы для интуиции» (2013)¹⁹. И в обоих случаях автор представлял лишь очень краткое разъяснение этого мысленного эксперимента. Поэтому я приведу собственную интерпретацию. С моей точки зрения, мысленный эксперимент Деннета иллюстрирует каузальную эффективность высокоуровневых свойств. Только этим можно объяснить функционирование этой сложной системы.

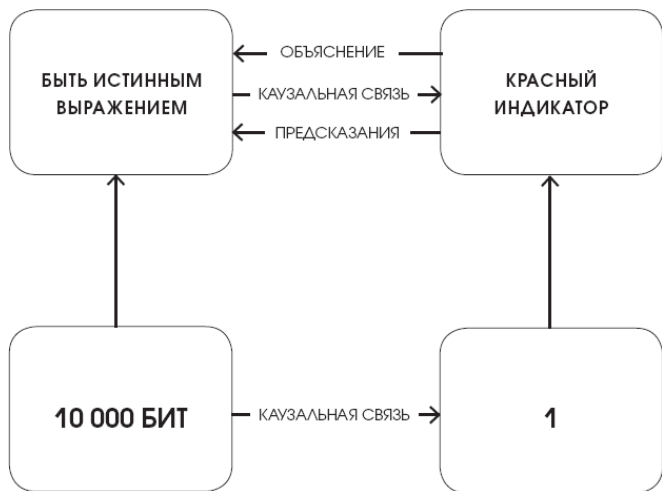


Рис. 3. Два черных ящика Д. Деннета

Описанная Деннетом система состоит из двух уровней. На

¹⁹ Также эта история была опубликована в журнале *Activitas Nervosa Superior* [Dennett 2010].

базовом уровне ее функционирование определяется последовательностями из символов: нулей и единиц. А на высоком уровне последовательностями из утверждений: утверждения могут быть ложными или истинными. Поскольку каждый раз из ящика *A* в ящик *B* передается новое закодированное сообщение, из одной только последовательности *символов* невозможно предсказать, какой индикатор загорится. Также с помощью одной только ссылки на последовательность входящих символов невозможно объяснить результат, цвет индикатора. На вопрос, почему в случае отправки некоторых последовательностей загорается красный индикатор, а не какой-то другой, есть только один удовлетворительный ответ: потому что эти последовательности в закодированном виде выражают истинные, а не ложные утверждения о мире. По замыслу Деннета, единственным способом объяснить и предсказать, почему и какой индикатор загорится, является ссылка на свойство высокого уровня, в данном случае – семантическое. Таким образом, результатом мысленного эксперимента можно считать следующие два тезиса:

(1) Механическая система, состоящая из двух соединенных ящиков и спроектированная для выполнения синтаксических операций, проявляет семантические свойства. То есть помимо низкоуровневых, базовых свойств, она обладает высокоуровневыми свойствами.

(2) Свойствами, которые определяют регулярность и каузальную связь между нажатием кнопок

на одном ящике и вспышками индикаторов на другом, являются эти семантические, высокоуровневые свойства последовательностей.

Синтаксические свойства, то есть последовательности битов, не содержат какого-то общего паттерна и потому не могут обосновывать каузальную закономерность. Синтаксис и семантика одновременно присутствуют в системе, но только семантические свойства имеют каузальную силу. Это не уникальный случай такого сочетания свойств. Вот, к примеру, человек идет по песчаному пляжу и оставляет за собой след. Этот след зависит от формы подошвы и веса человека, но никак не от цвета и материала подошвы. В предшествующем событии реализованы разные свойства, но только некоторые из них каузально эффективны. Вымышленная ситуация Деннета построена так, чтобы продемонстрировать, как семантическое содержание может быть каузально релевантным свойством, как форма или вес, а не как цвет или материал по отношению к следу на песке. Но, вероятно, каузальную действенность семантических свойств в механической системе можно было продемонстрировать как-то проще?

Аппарат Деннета напоминает светофор, управляемый вручную. Две кнопки на одном устройстве зажигают соответствующие лампочки на другом. Такая модель светофора демонстрировала бы (а) *каузальную регулярность*. Но этого недостаточно для философских целей. Для объяснения каузальной регулярности и предсказания состояний та-

кого «светофора» достаточно описаний физического уровня. Чтобы появилась (б) *возможность описывать работу «светофора» на интенциональном уровне*, система должна демонстрировать автономность, хотя бы в какой-то степени управляться изнутри и реагировать избирательно на входящие сигналы. Тогда можно будет описывать процесс так: отправленные первым компьютером данные второй компьютер «считает» соответствующими такому-то цвету. Кажется, необходимую автономность система начинает демонстрировать тогда, когда впервые загорается оранжевый индикатор. Но только возможности приписывать интенциональные состояния тоже недостаточно. Необходимо (в) *заблокировать возможность предсказаний на основе только лишь синтаксического анализа*. Для этого нужно создать ключ для генерации бесконечного числа последовательностей, сохранив при этом регулярность в работе системы. Ключ должен находиться вне системы, чтобы его нельзя было обнаружить и расшифровать. Таким ключом в аппарате Деннета является описание внешнего мира. «Именно тот факт, что каждый ящик несет в себе многообразную семантическую связь с одним объектом (описанием мира. – Д. В.), выраженную, однако, в разной “терминологии” и по-разному аксиоматизированную, является основой регулярности» [Dennett 2013, 195].

Возможно, ключом мог быть также какой-то внешний текст, например текст «Войны и мира» на двух языках. Но

тогда синтаксические последовательности, содержащие этот текст в закодированном виде, были бы конечными. А значит, со временем, даже не зная ключа (текста Толстого), ученые смогли бы предсказывать, какая последовательность будет «красной», а какая – «зеленой». Мне кажется, установка действительно соответствует минимальным требованиям и не чересчур усложнена. Но достигает ли Деннет с ее помощью поставленных целей? Удастся ли ему показать, что семантические высокоуровневые свойства способны обладать каузальной эффективностью в механических системах?

4. Критика телеофункционализма и возможные ответы

А. Невыполнимость условий мысленного эксперимента

На первый взгляд, описанный выше аппарат представим и возможен. Но уверенность в этом рассеивается, стоит лишь обдумать механизм его функционирования более тщательно. Видимо, программисты фактически все-таки не могли реализовать аппарат так, чтобы свойства «быть истинным» или «быть ложным высказыванием» регулярно вызывали соответствующую цветовую индикацию. Во-первых, истинность части высказываний должна со временем быть опро-

вергнута, во-вторых, что-то должно было измениться с момента создания базы и, в-третьих, то, что было истинным для машины *A*, могло не быть истинным для машины *B*. К примеру, высказывание «Изотоп антигелия-4 был зарегистрирован на ионном коллайдере RHIC» могло быть ложным до 2011 г., истинным после 2011 г. и даже бессмысленным до появления слов «изотоп», «антигелий» и «ионный коллайдер». А выражение «Мою программу написал Эл» истинно для первой машины и ложно для второй. Таким образом, в работе аппарата отсутствовали бы регулярности, которые описаны в мысленном эксперименте. И, при нажатии α , на ящике *B* иногда загорался бы красный индикатор, а иногда – зеленый и даже оранжевый.

Парировать возражение можно небольшим усовершенствованием эксперимента. Допустим, что программисты предварительно договорились, какие истины включить в экспертную программу. Тогда их базы данных должны соответствовать друг другу, а регулярность работы аппарата – соблюдаться. Но в таком случае неясно, что является каузально эффективным свойством. «Быть истинным» или «быть ложным выражением» в качестве таких свойств уже не работают. Так как в случае с предварительно согласованной базой высказываний программисты могли бы допустить совместные ошибки. Также в ходе исследования ящиков ученые могли случайно набрать последовательность символов, которая была бы истинной, но при этом не вошла бы в согласованный

массив знаний Эла и Бо. В качестве кандидата на каузальную эффективность можно принять свойство «быть истинным выражением для... машин А и В»²⁰. Но оправданно ли это?

В. Машина не обладает семантическими свойствами

Далеко не очевидно, что машины могут иметь убеждения. Для многих даже очевидно обратное. По крайней мере, машины, как они описаны в мысленном эксперименте Деннета, точно не представляются мыслящими существами. Поэтому эти машины вряд ли могут считать что-то истинным или ложным. Но высказывание может быть истинным или ложным для самих программистов. Тогда правильной всего свойством верхнего уровня считать только свойство «быть истинным или ложным для Эла и Бо». Иначе говоря, в аппарате Деннета нельзя найти высокоуровневых интенциональных свойств, для их наличия нужны программисты. Таким образом, мысленный эксперимент возвращает нас к исходным пунктам, к необъяснимой и противоречивой ментальной каузальности.

Так допустимо возражать, если предполагать возмож-

²⁰ Не факт, что Деннет согласился бы с такой поправкой. Возможно, она разрушает его построение. Ведь семантические характеристики происходят именно из связи с внешним миром. Они связаны с тем, что в последовательностях битов существует репрезентация мира, а не каких-то внутренних состояний.

ность каузальной связи между событиями, отстоящими во времени и пространстве. Ментальные состояния агентов, безусловно, имели отношение к работе этого аппарата тогда, когда программировалась база знаний. Но в момент нажатия кнопок они уже не имеют значения. Для верности можно представить, что два ящика были раскопаны следующими поколениями на месте, где раньше находился Лос-Анджелес, например в 3105 году. К тому времени программисты давно уже умерли. Вряд ли уместно говорить, что каузальным действием обладают ментальные состояния уже несуществующих людей.

На это можно ответить: пусть между ментальными состояниями программистов и состоянием компьютеров в момент исследования нет контакта, но генеалогическая связь прослеживается. Ментальные состояния программистов повлияли на исходные состояния компьютеров, а эти исходные состояния каузально связаны с каждым последующим их состоянием, и так – до самого последнего, до момента, когда ученые начали исследовать ящики *A* и *B* в лаборатории. Помогает ли такое рассуждение преодолеть временной зазор? Допустим, но тогда можно дальше усовершенствовать мысленный эксперимент. Например, можно представить, что массив истинных данных был получен путем сбора информации о внешнем мире непосредственно машинами-роботами. Тогда ментальные состояния не будут иметь прямого отношения к последовательностям битов. Впрочем,

скептик и здесь, наверное, не будет удовлетворен. Он может спросить: а реализуем ли такой сценарий? Возможно ли создать самообучающуюся энциклопедию? И даже если возможно, что будет обозначать информация, собранная машиной? Разве машина вообще может быть чувствительна к семантическим содержаниям? Она способна распознавать и совершать операции только с формами, содержание которых может быть интерпретировано сознательными мыслящими агентами. Но сама машина не способна интерпретировать эти сигналы. Только для мыслящих существ последовательность битов обозначает истинное высказывание или что-то еще. Только мыслящие существа наделяют материальные объекты смыслом. Машина – только инструмент, а никак не мыслящий агент.

Это возражение достаточно популярно. Одним из самых влиятельных сторонников такой позиции является уже известный нам и Деннету Джон Сёрл. Поэтому неудивительно, что в текстах Деннета можно найти развернутый подходящий ответ. Это – еще одна вымышленная история, или помпа для интуиции. Новая история Деннета должна подорвать уверенность в том, что только сознательные разумные существа наделяют знаки значением.

«Итак, предположим, вы решили по каким-либо причинам продлить свою жизнь и узнать, что произойдет в двадцать пятом веке» [Dennett 1996, 295], – начинает свою историю автор. Чтобы дождаться этого момента, придется дли-

тельное время провести в состоянии глубокого сна, комы. Когда вы проснетесь, ваших детей и внуков уже не будет. Так что нужно заранее позаботиться о сохранности тела, например создав специальную защитную капсулу. Создание такой капсулы для хранения – не единственная забота. Нужно обеспечить ее снабжение и безопасность. Выбирать придется между двумя вариантами. Первый – сразу найти идеальное место, куда будут подведены источники питания и энергии. Этот вариант плох тем, что обстоятельства могут измениться и идеальное место может перестать быть таковым. Второй вариант – создать движущуюся конструкцию, гигантского робота, и установить капсулу в него. Вторая, более сложная идея, требует, чтобы робот был способен «выбирать» то или иное действие исходя из ваших интересов. У него должны быть органы восприятия, способность управлять движением, даже планировать. В конечном итоге сложно предсказать, какие неприятности могут произойти в следующие столетия. К тому же нельзя исключать возможности создания других таких же роботов, которые будут бороться за ограниченные ресурсы, продолжает Д. Деннет. Вполне возможно, роботу придется состязаться или сотрудничать с ними. В таком случае его нужно запрограммировать так, чтобы, обладая основной целью (обеспечить ваше выживание), он мог выстраивать и свои, промежуточные. Вполне возможно, эти цели даже в какой-то момент станут преобладать, даже вступят в противоречие с основной целью. Его может толкнуть

на самоубийство другой робот, стремящийся подчинить все окружающее своей миссии. Или заставить помогать и защищать другой организм. В соответствии с описанием Деннета, очевидно, что программа и поведение такого робота должны быть очень сложными. Такой робот даже, наверное, прошел бы тест Тьюринга: по поведению его сложно будет отличить от обычного человека. Ему можно будет приписать подлинные интенциональные состояния. Он будет вести себя так, будто мыслит, имеет убеждения и желания... хотя все это будет запрограммировано вами заранее. Разве он не будет мыслящей машиной?

Если все это не убедило вас, значит, вне зависимости от деталей, вы не способны допустить интенциональность у любых механических объектов. Но если это так, вы... не способны допустить существование интенциональности даже у людей. История в руках Деннета вдруг обретает неожиданный оборот. До этого момента она могла показаться несуразной и уж слишком надуманной фантастикой. Но Деннет объявляет, что она – всего лишь немного замаскированная история эволюции человека, вариация на тему, представленную в знаменитой книге этолога и эволюционного биолога Ричарда Докинза «Эгоистичный ген» [Dennett 1996, 298].

Две стратегии выживания в биологической истории обозначают стратегии выживания двух царств: растений и животных. А сам робот – это мы. Как ни парадоксально, люди – это машины для выживания генов, продукт их творче-

ства. Это их эгоистичные планы мы выполняем, их цели, а не наши собственные, являются нашим *raison d'être*. В своей основе гены, конечно, весьма просты. Но они обладают бессознательной целенаправленностью: «желанием» сохранять собственную «жизнь» и преумножать число своих копий. И им удается реализовать свои планы. Жизнь одного гена намного дольше человеческой. Она почти бесконечна. Так кто кем в результате управляет, мы генами или гены нами? Кто стоит у руля? Правда, управление генов организмом происходит не в реальном времени, не как управление, скажем, марионеткой. Генам приходится предвидеть все заранее. Поскольку точные условия окружающей среды в будущем для них неизвестны, они строят самые общие предположения о рисках и возможностях. Направляя синтез белка, они программируют функциональность организма задолго до того, как «механизм выживания» встретится с реальными опасностями. Когда их работа проделана, организму приходится справляться самостоятельно. Но тем не менее именно гены остаются в центре эволюционного процесса и являются авторами «наших» замыслов.

Деннет считает, что наличие исходно семантической составляющей в нашем сознании – это иллюзия. «[Мозг] не может быть создан, чтобы решать невозможную задачу, но он может быть создан так, чтобы решать ее приближенно, *подражая* работе невозможного объекта (семантического процессора)» [Dennett 1996, 61]. Мозг решает эту задачу,

используя соответствия между структурными (синтаксическими) закономерностями в своем внутреннем устройстве и окружающей среде. Правда, что мы оперируем интенциональным содержанием, точнее, как бы оперируем. Но даже люди не обладают первичной, внутренней интенциональностью. Значениями состояния и действия людей наделяются постепенно, в ходе эволюционного процесса, во взаимодействии с миром. И этот процесс не имеет абсолютного завершения. Семантические свойства в любом случае производны. И четкой грани между только лишь «синтаксическими операциями» и «семантическими операциями» не может быть. Допущение четкой разделительной линии между ними – это тупиковый путь, ошибочный путь эссенциализма, на который философия все время норовит свернуть, может быть, еще со времен Сократа.

Абсурдность философского поиска сущностных свойств и проведения повсюду границ Деннет иллюстрирует с помощью псевдоаргумента, Аргумента о невозможности существования млекопитающих [Dennett 2013, 240]:

1. Каждое млекопитающее имеет мать-млекопитающее.
2. Если млекопитающие существуют, их к настоящему времени было конечное количество.
3. Но у каждого млекопитающего должна быть мать-млекопитающее. Тогда к настоящему времени должно быть бесконечное количество млекопитающих. Это

приводит к противоречию с пунктом 2.

Из этого аргумента будто бы следует, что млекопитающих не существует. Но это же абсурд. Несмотря на формальную корректность аргумента, следствие ложно. Ошибочны посылки, скрытое в них предположение наличия четкой грани между млекопитающими и их предками. И полностью автономные семантические двигатели – это такая же химера, как «первое млекопитающее». Млекопитающие постепенно эволюционируют из рептилий. И между двумя видами стоит множество транзитных «как бы млекопитающих» и их матерей (животных из отряда терапсидов). Аналогичным образом эволюционируют и семантические двигатели. Они не появляются внезапно. Более того, идеального семантического двигателя не существует. Поэтому вряд ли стоит удивляться, что воображаемый аппарат из двух ящиков не может абсолютно точно схватить свойство «быть истинным выражением». Он работает на «почти истинных выражениях» в большинстве случаев; впрочем, как и мозг человека.

С. Высокоуровневые свойства в принципе не обладают каузальной эффективностью

Впрочем, оппонент Деннета мог бы усомниться и в другом тезисе автора: в том, что в мысленном эксперименте с «Двумя черными ящиками» каузально эффективным свойством является какое-либо высокоуровневое свойство. Оп-

понент мог бы попытаться доказать, что в реальности каузальная сила есть только у другого, низкоуровневого физического свойства, например у свойства «быть частью такого-то блока данных». При реализации базы вероятней всего, что истинные выражения располагались в каком-то одном блоке, скорее всего даже пространственном, а ложные – в другом. Тогда каузальным действием обладают не семантические характеристики, а, например, пространственные.

Пространственное расположение действительно могло быть общностью, отличающей одни наборы битов от других. Но данные на жестком диске могли располагаться поразному. Даже если пространственная общность существовала, она не была обязательной. При смене «места» хранения последовательности она все равно вызывала бы ту же ответную реакцию второй машины. То есть пространственная общность являлась бы общностью случайной, в отличие от общности семантической. Это была бы случайная корреляция, а не каузальная зависимость.

Аналогичным образом в качестве претендента на каузальную роль не подходит и характеристика «принадлежать к блоку данных, связанных с α », так как вполне возможно, чтобы последовательности были связаны иначе, но регулярность бы соблюдалась. (Под этой регулярностью я подразумеваю отправку на второе устройство все время различных последовательностей битов без синтаксических и иных общностей и способность второго ящика отсортировать их

должным образом.) Для примера представим такое устройство, в котором металлический стержень соединяет между собой взрывчатое вещество и открытый огонь. Пока теплопроводность металлического стержня низкая, взрыва не происходит, так как недостаточно тепла передается от огня. Но стоит повысить теплопроводность стержня (например, заменив его на стержень с другим химическим составом), как произойдет взрыв.

Существует, однако, еще один верный способ произвести взрыв такого устройства. По закону Видемана – Франца, теплопроводность всегда при нормальных условиях коррелирует с электропроводностью (так как оба этих свойства обуславливаются движением свободных электронов). Соответственно, повышая электропроводность стержня в устройстве (заменяя химический состав стержня), мы также создаем условия для взрыва. При этом кажется, что электропроводность является только сопутствующим свойством, а не каузально эффективным²¹.

Описанное выше устройство демонстрирует, что только одно из двух коррелирующих между собой свойств может быть каузально эффективным. При этом в случае с данным устройством корреляция высокоуровневых свойств систематическая, регулярная. А в случае с пространственным расположением данных корреляция явно случайная, нерегу-

²¹ Подобный пример приводит Блок в статье «Может ли разум изменить мир?» [Block 1990, 147].

лярная, несистематическая. Последовательности битов могут располагаться как угодно, при этом истинные выражения, «выражения, истинные для машин Эла и Бо», будут вызывать один и тот же эффект. Контрфактический анализ помогает определить, какое из свойств антецедента является каузально эффективным, а какое – нет. И именно семантическое свойство выдерживает контрфактическую проверку, а не низкоуровневое физическое.

Д. Функциональные свойства не обладают каузальной эффективностью

Деннет считает, что семантические свойства систем могут быть каузально эффективны. При этом, как мы ранее установили, эти свойства, по его мнению, являются свойствами функциональными, то есть свойствами, которые определены через свой эффект. Но могут ли функциональные свойства быть каузально эффективными? Критиков этого утверждения не меньше, чем критиков интенциональности машин. Один из наиболее влиятельных среди них – философ из Нью-Йоркского университета, Нед Блок. В статье «Может ли разум изменить мир?» [Block 1990] он предлагает протестировать интуиции на нескольких примерах: свойствах «быть провокационным» и «быть снотворным».

Свойство «быть провокационным» определено через свой эффект – провокацию. Следовательно, оно является функ-

циональным. Провокационной можно считать, например, мулету тореадора, ведь она провоцирующим образом действует на быка – она возбуждает в нем агрессию. У мулеты в то же время, отмечает Блок, есть и другие свойства, в том числе свойство «быть красного цвета». Именно оно, по мнению Блока, провоцирует быка. Красный цвет каузально эффективен в отношении агрессивного поведения быка. Можно ли то же самое сказать в отношении «провокационности» мулеты? Провоцирует ли быка «провокационность» мулеты или только красный цвет?

«Провокационность» мулеты – это второпорядковое свойство. Как и другие функционально определенные свойства, оно зависит от наличия первопорядкового свойства. И именно свойство «быть красного цвета» находится в каузальных отношениях с событиями на входе и выходе. Видимо, в «провокационности» мулеты нет дополнительной каузальной работы. Бык недостаточно умен, чтобы осознать, что мулета провокационна. Ему достаточно красного цвета. Провокационность мулеты может спровоцировать членов Всемирного общества защиты животных, но не быка. Аналогично, считает Блок, можно рассуждать и о другом функциональном свойстве – «снотворности».

Это свойство тоже функциональное и тоже второпорядковое. Оно определено в терминах эффекта, результирующего состояния – сна. Но оно предполагает существование первопорядкового химического свойства, вызывающего сон.

«Быть снотворным» – значит обладать химическим свойством, каузально эффективным в отношении сна. Одни снотворные эффективны в силу наличия в них барбитуратов, другие – бензодиазепинов, но вряд ли какие-то – в силу свойства «снотворности». Только в случае, если пациент сам убеждает себя в «снотворности» лекарства, то есть в случае эффекта плацебо, лекарство может приводить к сну в силу самой «снотворности». Даже если химический состав медикамента каузально неэффективен в отношении сна, убеждение пациента в действенности препарата может быть «снотворным». Это, однако, особый случай. В целом вероятно, что свойство «быть снотворным» как функциональное свойство, как правило, каузально бездейственно.

По мнению Блока, примеры иллюстрируют тот факт, что функциональные свойства не могут быть каузально эффективными. Исключениями являются те особые случаи, когда они каузально эффективны в силу убеждений разумного существа. Но в стандартных случаях, то есть когда функциональные свойства определены посредством их результирующего эффекта и этот эффект не возникает в результате осознания разумного существа, они каузально бессильны [Block 1990, 156]. Блок считает, что интуиции в отношении этих примеров верны.

Эпифеноменальность функциональных свойств он объясняет тем, что функциональные свойства находятся в логической связи с эффектом. А для каузальной связи нуж-

ны контингентное, эмпирическое и номологическое отношения между событием и его следствием. Блок считает: медикамент обладает свойством «снотворности» потому, что обладает *другим* свойством, которое в случае употребления номологически гарантирует сон (или повышает его вероятность). То есть номологическое отношение имеет место между химическим или физическим свойством и эффектом, а не между эффектом и функциональным свойством. Когда мы объясняем нечто с помощью ссылки на функциональное свойство, мы прибегаем к аналитическому объяснению. Конечно, могут быть исключения, как в случаях с «провокационностью» мулеты для общества защиты животных или «снотворностью» плацебо. Но это – особые случаи, и они не соответствуют ситуации с семантическими свойствами. Это – мнение Блока, но у меня оно вызывает подозрения.

В рассуждениях Блокастораживает уже то, что в качестве примеров он использует сомнительные случаи и свойства. На самом деле красный цвет безразличен быку, так как у него дихроматическое зрение (впрочем, Блок это и сам не отрицает, и тем не менее – использует этот пример). Бык одинаковым образом реагировал бы на красную, синюю или зеленую ткань. Тогда что, если не красный цвет, возбуждает быка? «Провокационность» мулеты? Кажется, здесь Блок сам оказывается в ловушке собственной иллюстрации. Даже если мы согласны с его антифункционалистскими настроениями, этот пример не выполняет своей функции. Со сно-

творным качеством пример получается тоже не очень удачным. Блок как бы нарочно напоминает о псевдообъяснениях. Вспомните мольеровского «Мнимого больного»: опиум действует снотворно, так как в нем имеется «снотворная сила». Мольер использует это объяснение с сарказмом, высмеивая неосведомленность объясняющего. Возможно, на эти ассоциации частично опирается и Блок. Но это – психологический прием, а не убедительное рассуждение. Вряд ли сторонник функционализма сам стал бы использовать «снотворность» лекарства и «провокационность» мулеты в качестве иллюстраций каузально эффективных функциональных свойств. В этом конкретном случае Блок борется с ветряными мельницами. Но кроме неудачных примеров в его аргументации можно обнаружить уже знакомый нам тип ошибки. Эта ошибка близка той, что допускали Э. Энском и Г. Райл (а кроме них – также А. Мелден и др.), когда критиковали действенность ментальных состояний с помощью Аргумента логической связи.

Тогда, напомню, речь шла о ментальных состояниях безотносительно их функционалистской интерпретации. И тем не менее рассуждения Блока напоминают этот аргумент. Если логика Блока, Энском и Райла схожа, то против них можно выстроить защиту наподобие той, что выстраивалась в ответ на АЛС. Можно попытаться разъединить строгую концептуальную связь, например, показав, что наличие того или иного функционального свойства не обязательно приво-

дит к ожидаемому эффекту. Функциональные свойства сосуществуют и взаимодействуют с другими функциональными свойствами, и результаты этих взаимодействий могут давать различный эффект. Одни функциональные свойства каузально эффективны, а другие – нет. Если бы существовала жесткая логическая сцепка между всеми функциональными свойствами и их эффектами, таких вариаций не было бы.

Возьмем свойство «быть будильником». Это – функциональное свойство, так как оно определяется в терминах функции, роли, которую играет это устройство, а не материала или его внутреннего механизма. Будильник – это устройство (электронное, механическое или даже органическое!), в заданный момент подающее сигнал к пробуждению, эффект которого – пробуждение. Допустим, некто пытается дать такое каузальное объяснение: он проснулся в шесть утра от звонка будильника. По логике Блока, раз будильник будит, то логически следует, что он должен разбудить. Но разве из того, что будильник будит, следует то, что звонок раздастся в шесть утра? Звонок мог раздаться в любое время, причем даже не в то время, на которое был поставлен: часы могли отстать или заторопиться, да и будильник мог вообще сломаться или быть незаведенным. К тому же тот, кого будят, мог и не проснуться от его звонка, несмотря на то что функциональное свойство будильника – будить. «Быть будильником» и «будить», безусловно, как-то логически состыкованы, но без номологической связи одно все равно не будет

причиной другого.

Ранее я давал функциональное определение боли: состояние, вызванное повреждением органов, сопровождающееся «убеждением, что что-то повреждено», «желанием это прекратить» и ведущее к реакции избегания. Но разве обязательно, чтобы испытывающий боль демонстрировал «реакцию избегания»? Если человек, к примеру, под пытками стремится убедить своих палачей, что он не способен чувствовать боль, не стонет, не выражает страданий в мимике, разве он не испытывает боль? Или если перед быком тореадор размахивает провоцирующей мулетой, но бык так и остается бездвижным (это ведь вполне возможно) – что, мулета перестает быть провоцирующей или перестает быть красной? Нет, просто бык мог устать. Ментальные состояния, определенные функционально, всегда взаимодействуют с другими ментальными состояниями (убеждениями и желаниями) и могут не образовывать жестких логических связей с последующим действием. Это значит, что семантические свойства в функциональной интерпретации, вероятно, могут быть каузально эффективными. Впрочем, не думаю, что эти соображения убедили бы всех противников ментальной каузальности, в частности Дж. Кима. Необходимо ответить на ключевой аргумент оппонентов – Аргумент исключения. Для борьбы с ним я хочу воспользоваться более сложной тактикой.

Решение будет состоять из трех шагов. На первом шаге

я представлю аргумент, аналогичный Аргументу исключения, но ведущий к очевидно ложным следствиям. Этот аргумент, как это ни удивительно, разработал уже известный нам Н. Блок. Вторым шагом я приведу иллюстрацию каузальной эффективности высокоуровневых свойств и представлю ее анализ. На третьем шаге я покажу ошибки в механике Аргумента исключения Дж. Кима. Это трехчастное решение проблемы ментальной каузальности работает в рамках телеофункционалистской теории сознания, которую я отстаиваю.

VI. Решение проблемы ментальной каузальности

1. Аргумент обобщения Н. Блока против аргумента исключения

Зачастую лучшая тактика защиты – это нападение. И в данном случае такая тактика уместна. Можно показать, что Аргумент исключения, направленный против ментальной каузальности и нередуцируемых свойств, противоречит важным научным утверждениям. Он подвергает сомнению установки научного знания, в частности убеждение в каузальной эффективности свойств и объектов, входящих в теории специальных наук. Под специальными науками имеются в виду все науки, отличающиеся от фундаментальной физики. Чтобы показать это, необходимо найти сходный по логической структуре и опирающийся на те же ключевые посыпки аргумент-двойник. Но он должен быть направлен не против ментальной каузальности, а против чего-то, что оспорить значительно труднее. Подобный метод аргументации был уже мною использован выше для критики Аргумента каузальных траекторий. В данном случае он тоже применим. В качестве аргумента-двойника можно использовать уже разрабо-

танный Аргумент обобщения, предложенный упомянутым нами Н. Блоком [Block 2003]. Аргумент называется так потому, что обобщает, экстраполирует выводы, сделанные Кимом относительно эффективности ментальных свойств, на эффективность свойств, входящих в аппарат специальных наук. Суть рассуждения – в следующем.

Допустим, что все свойства макрообъектов определены свойствами их частей, вплоть до элементарных частиц. Тогда свойства макрообъектов супервентны на свойствах микрообъектов. При этом свойства микрообъектов являются достаточной причиной для всех событий с макрообъектами. Тогда у свойств макрообъектов нет каузальной роли и, следовательно, согласно Принципу исключения, эти свойства эпифеноменальны (то есть каузально неэффективны). Они проигрывают в конкуренции за каузальную эффективность. Эта логика применима и ко второпорядковым свойствам. Первостепенными свойствами являются свойства индивидов. Второстепенные свойства – свойства свойств первого порядка. Все второстепенные свойства супервентны на первостепенных свойствах. Если Аргумент исключения верен, то и второстепенные свойства не играют каузальной роли. Но такими являются все свойства объектов специальных наук. Следовательно, химические, биологические и физиологические свойства тоже не являются каузально действенными. Все объекты и их свойства за пределами физики являются эпифеноменальными, а законы специальных наук

не схватывают каузальных регулярностей. Иными словами, «все науки делятся на физику и коллекционирование марок». Такой вывод, кажется, следует из Аргумента обобщения.

А что, если элементарные частицы физики делимы до бесконечности и фундаментального, базового уровня вообще не существует? На каком уровне тогда искать каузальные закономерности? А ведь это – реальная возможность, обсуждаемая современной физикой. Тогда проблема представляется еще более драматичной. Каузально эффективными перестают быть все известные нам объекты, и даже открытые в будущем объекты будут каузально бездейственными. Теряется вообще надежда открыть каузальные законы. Поиск таких законов становится подобным заполнению водой бочки с пустым дном. С открытием каждого нового, более фундаментального уровня каузальность перетекает в него, и так до бесконечности. Выявление причин превращается в погоню за неуловимой целью. Кажется, перед нами новая дилемма: принимать Аргумент исключения и вместе с ним – Аргумент обобщения, требующий редукции всех специальных наук (да и значительной части физики), или отрицать Аргумент исключения.

В пользу второй альтернативы, отрицания АИ, кажется, достаточно доводов. С практической точки зрения специальные науки вполне оправдывают себя. Они работают с регулярностями, на которые можно положиться. К тому же до-

водом в пользу отрицания АИ является и ментальная каузальность. Если АИ поражает одновременно все супервентные свойства, ментальные свойства оказываются «в хорошей компании». А проблема ментальной каузальности становится частным случаем проблемы всех высокоуровневых наук. Это значит, что проблемы вообще может не быть. Если следствие, к которому ведет аргумент, абсурдно, нужно проверить посылки или форму аргумента. Не наш ли это случай? Быть может, здесь заложена какая-то ошибка, и Аргумент обобщения сводит к абсурду Аргумент исключения?

Такого казуса можно попытаться избежать. Автор Аргумента исключения Дж. Ким считает, что существует путь между Сциллой и Харибдой. Можно сохранить АИ против нередуцируемых ментальных свойств и не попасть в ловушку Аргумента обобщения. Для этого нужно показать, что утечка каузальной эффективности от высокопорядковых к низкопорядковым²² свойствам не ведет к утечке каузальной эффективности от макро- к микрообъектам. Свойства объектов, рассуждает он, действительно образуют особую иерархию. Над свойствами нижнего порядка надстраиваются свойства верхнего: над физическими – химические, над химическими – биологические и т. д. Ментальные свойства находятся где-то в верхней части иерархии. Они надстрое-

²² Здесь я вслед за Кимом нарочно использую термины «высокопорядковое» и «низкопорядковое» свойство, чтобы зарезервировать термин «уровень» исключительно за иерархией макро-микро.

ны над биологическими, так как все ментальные свойства присущи только системам с биологическими свойствами (по крайней мере, до создания ИИ) и не наоборот. Но вся эта иерархия свойств не является иерархией от малого к большому. Она построена не движением от микро- к макрообъектам, а с помощью отношения реализации: физические свойства реализуют химические свойства, химические – биологические. По шкале микро-макро движения нет. Биологическими, ментальными и физическими свойствами обладает один и тот же объект. Таким образом, регресс от высокопорядковых свойств к низкопорядковым происходит внутри одного и того же класса объектов. К примеру, свойством «весить 80 кг» может обладать объект, обладающий свойством «быть *Homo sapiens*» и свойством «желать горячего чая». Разные свойства, свойства разного порядка, относятся к одному объекту.

Таким образом, из Аргумента обобщения не следует бесконечный регресс в микро-макро иерархии. Макрообъекты обладают своими особыми каузальными свойствами, отличными от свойств составляющих их частей. Следовательно, атомы и составляющие их частицы не имеют абсолютной монополии на каузальную эффективность. И возможность деления этих частиц до бесконечности не представляет угрозы для каузальности в целом. Впрочем, так Ким блокирует только один вариант утечки каузальности: от макро- к микрообъектам. Но утечка может проходить и в другом направ-

лении: от высокопорядковых свойств к первопорядковым.

Пусть составные объекты обладают каузальной силой, отличной от каузальной силы их частей, но возможно, только самые фундаментальные их свойства каузально эффективны? Утечка каузальности может произойти не вниз, с уровня макро- на уровень микрообъектов, а как бы «в сторону», на одном уровне. Если только фундаментальные физические свойства объектов каузально эффективны, свойства, которые изучают специальные науки, оказываются эпи-феноменальными. Тогда законы этих наук не являются каузальными законами. Впрочем, остается ключевой вопрос: какие свойства должны попасть в категорию фундаментальных физических? Только масса, спин, электрический заряд? А должны ли, например, свойства «быть жидким» или «быть газообразным» попасть в эту категорию?

Кажется, с одними только базовыми свойствами не объяснить растворяющее действие жидкостей или процесс кипения. Для того чтобы физический мир был каузально замкнутым, чтобы имеющиеся в нем феномены можно было полностью объяснять, требуется значительно расширять набор базовых физических свойств. И в него, по-видимому, войдут свойства как элементарных частиц, так и атомов, молекул, клеток, сложных живых организмов, планет, звезд, то есть свойства, приписываемые объектам в химии, биологии, физиологии, геологии т. п. Среди них будут и функциональные свойства. При этом такие свойства не будут каузально неэф-

фективными, а только «каузально гетерогенными». То есть их действенность будет определяться их реализацией в других свойствах.

Разъяснения, представленные выше, отчасти спасают АИ от абсурдных следствий. В частности, они удостоверяют, что АИ не ставит под сомнение каузальные свойства макрообъектов. Но могут остаться сомнения в отношении следствий для специальных наук. Действительно ли из АИ не следует, что свойства объектов, приписываемые им в специальных науках, эпифеноменальны? Действительно ли сеть, образуемая АИ, захватывает только нередуцируемое ментальное содержание? Защита Дж. Кима построена на представлении о свойствах с микробазой. Именно опора на это представление позволяет философу включить в физический каузально эффективный мир высокопорядковые свойства. Свойства с микробазой, по его мнению, – это любые свойства, у которых в качестве основания имеются микросвойства. «Свойство с микробазой некоторого объекта – это свойство, характеризующее его микроструктуру. Оно говорит нам о том, из какого рода микроскопических элементов состоит этот объект, и о структурных отношениях, которые формируют эти элементы в устойчивый объект» [Kim 2008, 57]. По мнению Кима, все специальные науки оперируют именно такими свойствами.

Вот свойства молекул воды, например, имеют микробазу, так как основаны на свойствах атомов водорода и кислоро-

да. Свойства воды редуцируемы к свойствам H_2O , ведь вода, считает Ким, и молекулы H_2O редуцируемы друг к другу. И именно поэтому свойства воды не конкурируют со свойствами водорода и кислорода за каузальную эффективность. Ведь нечто не может конкурировать с самим собой. Если каузально эффективна основа, каузально эффективны и производные свойства. Так, психологические свойства каузально эффективны в силу своей реализации и тождественности со свойствами биологическими, биологические – с химическими и т. д. Видимо, это Ким и имеет в виду, когда говорит о каузальной гетерогенности свойств высокого порядка. С ментальными свойствами (по крайней мере, качественными), однако, ситуация другая. В отличие от свойств, приписываемых объектам в специальных науках, ментальные свойства не имеют микробазы, они нередуцируемы и поэтому выступают конкурентами физических в борьбе за каузальную значимость. Соответственно (по АИ), в этой борьбе они проигрывают.

Впрочем, в такой позиции есть еще брешь. Не только ментальные качественные свойства сопротивляются редукции. Вызывает сомнение возможность редукции высокопорядковых свойств в специальных науках к свойствам фундаментальной физики. И камнем преткновения здесь является *множественная реализация*. Некоторые высокопорядковые свойства могут иметь различную реализацию на свойствах более низкого порядка. В высокопорядковых науках

они могут рассматриваться как один естественный вид, а в низкопорядковых – как множество. Вот даже вода: ее тождественность с молекулами H_2O – уже слишком грубое обобщение. Фактически проба воды дает не только мономерные молекулы H_2O , но и полимерные, такие как $(\text{H}_2\text{O})_2$ и $(\text{H}_2\text{O})_3$. Все разновидности молекул находятся в статистическом балансе, но постоянно образуют короткоживущие связи. То есть даже вода может иметь множественную реализацию [Block 2003]. Более того, как свидетельствуют некоторые химики, даже сама процедура идентификации воды состоит не в анализе составляющих ее элементов, а в тестировании ее функциональных свойств. Определение естественного вида в химии может проходить не с помощью анализа составляющих, а посредством анализа реакций. Значит, многие важнейшие свойства, изучаемые специальными науками, – свойства функциональные.

Примеров естественных видов и высокопорядковых свойств с множественной реализацией достаточно. Так, даже свойство «быть твердым телом» имеет как минимум две реализации: в кристаллической решетке (как у алмаза) и в аморфной субстанции (как у стекла). В кристаллах атомы расположены регулярно, с периодичностью. А в аморфных телах они колеблются вокруг хаотически расположенных точек. Но тем не менее и аморфные, и кристаллические твердые тела представляют собой один естественный вид, реализуют одно свойство. Они подчиняются единообразным зако-

нам, которые изучает подраздел физики – механика твердого тела.

Все эти соображения показывают, что естественные виды, фигурирующие в законах специальных наук, могут быть по-разному реализованы. А следовательно, они не редуцируются к одному свойству как к микробазе. А если они не редуцируются, значит, они конкурируют за каузальную эффективность с базовыми первопорядковыми физическими свойствами и в соответствии с Принципом каузальной замкнутости физического мира должны проигрывать в этой конкуренции и быть эпифеноменальными. Таким образом, Аргумент исключения исключает из числа каузально эффективных свойств множество свойств, составляющих теоретическую базу специальных наук, подрывает статус их законов и тем самым сам сводится к абсурду. Ибо такое следствие вносит слишком много разрушений и, соответственно, неприемлемо. Все эти рассуждения приводят нас к мысли, что, если ментальная каузальность и выглядит на каком-то основании проблематичной, она представляется не более проблематичной, чем каузальность второпорядковых свойств и свойств, изучаемых специальными науками. То есть ментальная каузальность находится «в хорошей компании». Это – первый шаг критики Аргумента исключения.

Итак, можно критиковать Аргумент исключения с помощью аргумента-двойника, такого как Аргумент обобщения, и его следствий. Следующее направление критики – с помо-

щью контрпримеров. Можно найти или придумать ситуации, в которых причинами чего-либо будут выступать не фундаментальные физические, а высокопорядковые свойства. Для критики аргумента Кима здесь я хочу предложить собственный пример, который должен продемонстрировать каузальную работу «исторических» свойств. Исторические свойства объектов не являются базовыми физическими свойствами. Если они имеют каузальное влияние, то оно может быть и у ментальных событий. Замысел предлагаемой мной истории сходен с замыслом автора «Истории про два черных ящика», но есть и важное отличие. Новый пример позволяет продемонстрировать, как одинаковые по фундаментальным физическим свойствам события благодаря свойствам высокого порядка могут приводить к разным действиям. Это мне представляется преимуществом нового мысленного эксперимента.

2. Контрпример к аргументу исключения. Библиотека первых изданий

Представим, что в Марракеше живет старый библиотекарь, который посвятил свою жизнь коллекционированию древних философских книг. Он собирает первые издания философских трактатов. В его библиотеке хранятся почти все первые издания древних философских манускриптов,

папирусов, сочинений. Библиотекарь знает историю каждой философской книги в мире и может датировать абсолютно любую работу. Поэтому в его библиотеке нет ни одного произведения, которое было бы не первым изданием. К этому вопросу он относится особенно щепетильно. И вот в его день рождения на его столе оказываются две совершенно одинаковые книги. Это – отсутствующее в его коллекции произведение, подарок от его друга, коллекционера из Рабата. Книги абсолютно идентичны. Это – полные физические дубликаты. Увидев подарок, старый библиотекарь улыбается, внимательно их рассматривает, а потом одну книгу ставит на библиотечную полку, а другую требует отослать своему знакомому обратно в Рабат. Как объяснить такое странное поведение – ведь библиотекарь собирает *все* первые издания философских книг?

Дело в том, что библиотекарь знает каузальную историю абсолютно каждой книги. И знает, что по заказу его друга была сделана абсолютно точная физическая копия первого издания этой книги. Таким образом, одна книга является первым изданием, а другая – нет. Именно поэтому только один из томов подходит для его коллекции.

В вымышленной ситуации одна книга – в библиотеке древних изданий в Марракеше, а другая отправлена в Рабат. При этом книги физически абсолютно идентичны. Как объяснить разную траекторию этих книг? Из описания ситуации разница между книгами понятна: она в том, когда и как были

изданы эти экземпляры. Только одна книга оказалась в библиотеке потому, что только она была первым изданием. Это объяснение вполне удовлетворительно. Оно позволяет предсказать развитие аналогичных событий и в будущем. Если вдруг в следующий раз старому библиотекарю придут еще две книги, мы уже не так удивимся его поведению. И при наличии абсолютно полных знаний сможем в другой раз сами предсказать, какая книга окажется на полке, а какая – на почте. Но какого рода это объяснение? Я считаю, что это объяснение – каузальное. Оно выявляет причину различной траектории книг.

О сущности каузальных связей у философов нет согласия. Существует множество теорий относительно того, что лежит в основе суждений о каузальности. В этой работе нет возможности обсуждать детали этих разногласий. Однако существуют общие места, точки согласия, самые общие критерии каузальности, которые обычно принимаются. На них мы и будем опираться. Этими общими критериями каузальной связи являются регулярность, контрфактическая зависимость, временное и пространственное соседство. Итак, проверим соответствие нашего описания этим критериям.

Во-первых, мы видим присутствие регулярности. Событие «поступление первого издания философской книги» регулярно сочетается с событием «пополнение библиотеки». В этой регулярности нет и не предвидится исключений. Кто-то может возразить: но библиотекарь, в конце концов, мо-

жет ошибиться или поменять свои критерии. Тогда регулярность нарушится! Это означает, что регулярность – простое совпадение. Что ж, тогда можно усовершенствовать мысленный эксперимент: вместо человека-библиотекаря поставить машину – робота, который будет обладать всей информацией и отслеживать каузальную траекторию всех книг. Такой робот будет механически отсортировать древние книги от подделок и сохранит требуемую регулярность. Таким образом в мысленном эксперименте будет выполнено первое требование.

Во-вторых, между событием «поступление первого издания философской книги» и событием «пополнение библиотеки» существует контрфактическая зависимость. Если произойдет событие «поступление первого издания философской книги», то произойдет событие «пополнение библиотеки», и если событие «поступление первого издания философской книги» не произойдет, то не произойдет и «пополнение библиотеки». Таким образом, выполняется и второе требование к каузальной связи. К этому можно добавить и другие традиционные критерии: предшествование во времени причины действию и их пространственную смежность. Эти критерии тоже выполнимы. Кажется, можно заключить, что связь между событиями каузальная. Теперь стоит уточнить, благодаря каким именно свойствам осуществляется эта каузальная связь.

Книги являются физическими объектами. У них есть фи-

зические свойства. И они могли бы выступать носителями каузальных закономерностей. Но в предложенном мной случае они не имеют каузальной эффективности. Дело в том, что книги, по условиям эксперимента, являются физическими двойниками, то есть в физическом отношении ничем не отличаются друг от друга. И тем не менее они оказываются в разных местах: одна – в Марракеше, а другая – в Рабате. Причиной различий их траекторий служит не какое-то их физическое свойство, а каузальная история, свойство «быть первым изданием философского трактата» или, что более точно, реляционное отношение между конкретным экземпляром книги, ее местонахождением и воспоминаниями библиотекаря или состояниями блока памяти робота, отношение R . Это реляционное свойство в данной интерпретации каузально действенно. Следовательно, это – контрпример для аргумента Кима. Чтобы убедиться в этом, рассмотрим три наиболее вероятных возражения.

Возражение 1. Можно было бы показать каузальную эффективность высокопорядковых свойств на более тривиальном примере

Для того чтобы привести контрпример аргументу Кима, большинство тривиальных случаев не подходит. Так, во множестве стандартных ситуаций высокопорядковые свойства реализованы локально в самом объекте. Интерпретируя эти ситуации, можно настаивать на утечке каузальности к низкопорядковым физическим свойствам. В нашем же случае

локальные физические свойства объекта недостаточны для объяснения устойчивой закономерности. Несмотря на то что книги неразличимы по своим физическим параметрам, их появление приводит к различным действиям. Для объяснения этого необходимо соотношение каузальной траектории книги и памяти библиотекаря или робота. Более того, в описанной ситуации могут возникать любые физические события (могут поступать любые книги), при этом закономерность может быть каузально объяснена только через обращение к высокопорядковому реляционному свойству R .

Возражение 2. Аргумент спасает каузальную эффективность высокопорядковых свойств, но не спасает ментальную каузальность

Предположим, что интерпретация описанной выше ситуации верна. Но разве она позволяет нам утверждать, что ментальные свойства каузально эффективны? Свойство «быть древним философским трактатом» – это просто реляционное отношение между экземпляром книги (конкретным физическим объектом), ее местоположением и информацией в памяти библиотекаря или блоке памяти робота (также реализованной в физическом носителе). Это свойство R – просто отношение между физическими объектами. В свою очередь, ментальные состояния таковыми не являются. Ментальные состояния – это субъективные качественные состояния. Доказав каузальную эффективность свойств типа R , мы еще не доказали возможность каузальной эффективно-

сти ментальных состояний. Так мог бы рассуждать оппонент нашего аргумента.

Но мне эта критика не представляется основательной. Целью нашего эксперимента было показать, что каузально эффективными могут быть любые другие свойства помимо базовых физических свойств. Свойство R не является базовым физическим свойством, хотя оно глобально супервентно на этих свойствах. Поэтому наша цель достигнута. Возможно, что и ментальные свойства являются подобными высокопорядковыми свойствами и что они супервентны, хотя и нередуцируемы к базовым свойствам физического субстрата.

Возражение 3. Аргумент не спасает каузальную эффективность высокопорядковых свойств

Состояния библиотекаря (или робота) различны в случаях предъявления ему различных по истории, но идентичных физически книг. Эти различные состояния и являются причиной разных действий, – мог бы возразить оппонент. С этим возражением я, безусловно, согласен. Но вопрос именно в том, почему библиотекарь приходит к таким разным состояниям, ведь перед ним появляются физически идентичные книги. Для библиотекаря они, конечно, с самого начала не идентичны: он знает, какая из них старая, а какая – нет. Но именно это и показывает, что у книг имеются каузально эффективные высокопорядковые свойства. «Старая» и «новая» не есть свойства фундаментальной физики. При том что только эти свойства дают единственно доступную

возможность объяснить каузальную закономерность.

Оппонент мог бы сказать, что тогда можно упростить условие мысленного эксперимента. Например, представить, что библиотекарь всегда берет книгу, которая справа, и отправляет обратно книгу, которая слева. В таком случае реализация одинаковых физических событий тоже вела бы к разным действиям. Но при этом вряд ли можно было бы говорить о каузальной эффективности высокопорядковых свойств.

На это можно, в свою очередь, возразить: книги могли бы оказаться по разные стороны от библиотекаря. И тогда последовали бы другие действия. Пример с таким алгоритмом не схватывает интересной «высокопорядковой» закономерности. В то время как наш пример с Библиотекой первых изданий философских трактатов схватывает важную каузальную закономерность, которая построена на высокопорядковых свойствах книг. Какой бы физический объект ни оказался перед библиотекарем, с какой бы стороны он ни оказался, библиотекарь поступит с ним в соответствии с его каузальной историей.

3. Что не так с аргументом исключения Дж. Кима?

Аргумент обобщения доказывает, что если ментальная каузальность и проблематична, по такой же логике пробле-

матична каузальность всех свойств, приписываемых объектам в специальных науках, то есть во всех науках, кроме фундаментальной физики. Примеры с Двумя черными ящиками и Библиотекой первых изданий показывают, что высокопорядковые свойства иногда являются единственным способом объяснить регулярности в мире. Это указывает на их каузальную эффективность. Но для решения проблемы остается выяснить, что же тогда «не так» с аргументом Дж. Кима. Чтобы это показать, удобно представить аргумент в формальном виде.

(ПКЗ¹) *Если физическое событие имеет причину в момент t , оно имеет достаточную физическую причину в момент t .*

(ПКИ) *Если событие e имеет достаточную причину c в момент t , никакое событие в t , отличающееся от c , не может быть причиной e (кроме случаев сверхдетерминации).*

(ПОСС) *Не существует систематической сверхдетерминации физических событий.*

Структура рассуждения Кима такова:

1. Допустим, ментальное событие m^0 является причиной физического события p , $m^0 \rightarrow p$.
2. Согласно ПКЗ¹, у события p есть также достаточная физическая причина p^0 , $p^0 \rightarrow p$. (ПКЗ¹)

3. $m^0 \neq p^0$.

4. Это – не случай свехдетерминации. (ПОСС)

5. Должна быть одна достаточная причина. (ПКИ)

6. Этой причиной является физическая причина p^0 .

(ПКЗ¹)

Следовательно, у ментального события нет каузальной действительности. Вывод: m^0 – эпифеноменально.

Думаю, что в описанном выше рассуждении ошибка заключается в применении ПКЗ¹. ПКЗ¹ допускает широкую и узкую трактовку физических свойств. Если использовать узкую трактовку, то получается, что причинами всех событий должны быть фундаментальные физические свойства. И эти свойства конкурируют за каузальную эффективность с ментальными. Но это не обязательно должно быть так. Некоторые регулярности существуют именно из-за высокопорядковых свойств. Эти свойства подпадают под широкую трактовку физических свойств. Ментальное как высокопорядковая характеристика физического может быть частным случаем физических свойств в широком смысле. Тогда будет неправильно полагать, что, помимо высокопорядковых свойств, причиной являются еще какие-то свойства. То есть шаг 2 в рассуждении окажется ложным.

Одна из последних работ Кима называется «Физикализм или что-то около того». Это – наиболее актуальное и полное изложение его метафизической позиции. Как следует из на-

звания, позиция автора – это «почти физикализм». Судя по аргументу, мне, однако, кажется, что позиция автора – это не «почти физикализм», а даже «нечто большее, чем физикализм». Это – некий «пуританский физикализм». Такой физикализм разделять мы вовсе не обязаны, тем более в ущерб ментальной каузальности.

VII. Итоги первой главы

В начале главы мы выяснили, что ментальная каузальность является необходимым условием свободного и ответственного действия. Однако она оказывается проблематичной в свете аргументов, предложенных современными метафизиками. Основным препятствием служит Аргумент исключения Дж. Кима. Он построен на достоверных посылах, в частности на принципе физикализма, и тем не менее ведет к отрицанию ментальной каузальности.

Существует несколько вариантов защиты ментальной каузальности. В первую очередь это локальный интеракционизм. Эта теория во многом опирается на Аргумент каузальных траекторий российского философа В. Васильева. Аргумент направлен на то, чтобы показать ложность Тезиса о локальной супервентности и обосновать особую роль ментальных состояний в поведении физических объектов. Однако это рассуждение не лишено недостатков и опровергается с помощью предложенного мной Аргумента-двойника.

Другой вариант защиты ментальной каузальности представлен в биологическом натурализме. В рамках этой теории ментальная каузальность вовсе не является проблематичной. Однако сам этот подход внутренне противоречив. Его главный сторонник, американский философ Дж. Сёрл, смешивает понятия реализации, супервентности и каузаль-

ности. Это приводит к новым парадоксам. Из проведенного анализа следует, что решение проблемы ментальной каузальности в теории биологического натурализма не является оптимальным.

В качестве третьей альтернативы мы рассмотрели теорию другого американского аналитика, Д. Деннета, – телеофункционализм. Деннет убежден, что проблема ментальной каузальности решена. Его позиция представлена в виде мысленного эксперимента – ситуации, которая должна иллюстрировать каузальную эффективность высокопорядковых свойств. Однако сама по себе эта иллюстрация недостаточна для того, чтобы служить решением проблемы ментальной каузальности. Отсутствие у Деннета разъяснений не позволяет сделать вывод о том, что именно в Аргументе исключения он отрицает. Ввиду этого я предлагаю трехчастное решение.

Это решение основано на Аргументе обобщения, предложенном Н. Блоком, контрпримере библиотеки первых изданий и критике Аргумента исключения. Предложенный подход снимает проблему сочетания ментальной каузальности и физикализма и обосновывает возможность действия. Теперь, когда одно препятствие на пути к свободе воли и моральной ответственности преодолено, мы можем перейти к следующему – детерминизму.

Глава 2. Свобода воли и детерминизм

I. Проблема соотношения свободы, детерминизма и индетерминизма

Предыдущую главу я посвятил анализу первого препятствия на пути реализации свободы воли – физикализма. Во второй главе я хочу обратить внимание на другое потенциальное препятствие – детерминизм. В самом начале главы я расскажу о связи понятий свободы и ответственности. Я объясню, почему понятие ответственности является центральным в дискуссиях о свободе воли. Потом я покажу, как понятие ответственности позволяет редуцировать свободу воли к двум базовым принципам: Принципу альтернативных возможностей и Принципу автономии (или «ответственности в конечном счете»). Далее я познакомлю читателя с Аргументом последствий и Аргументом манипуляций. Эти аргументы как раз показывают противоречие между двумя базовыми принципами и детерминизмом и тем самым выявляют проблему свободы воли.

В качестве возможных способов решения проблемы свободы воли я сначала представлю две наиболее влиятельные

современные инкомпатибилистские позиции: либертарианство Р. Кейна и жесткий инкомпатибилизм Д. Перебума. Обе эти позиции отрицают совместимость детерминизма, свободы и ответственности. Согласно либертарианству Кейна, свобода и ответственность несовместимы только с детерминизмом. Основанием его теории является Аргумент ответственности в конечном счете. Согласно позиции Перебума, свобода и ответственность несовместимы как с детерминизмом, так и с индетерминизмом. Основанием его убеждений является Аргумент исчезающего агента. Однако обе эти позиции, несмотря на аргументы в их пользу, по моему мнению, неудовлетворительны.

Дальше я представлю критику инкомпатибилизма. Я анализирую примеры, предложенные Г. Франкфуртом в 1969 г. против Принципа альтернативных возможностей, суждения Д. Деннета против Аргумента последствий, а также предложу критику Аргумента манипуляций Д. Перебума и собственный ответ на Аргумент исчезающего агента – Аргумент четырехмерного агента. Эти рассуждения станут основанием позиции суперкомпатибилизма – утверждения о том, что свобода и ответственность совместимы как с детерминизмом, так и с индетерминизмом. Эту позицию я считаю наиболее достоверной.

А. Ответственность как индикатор свободы

Условием свободы является отсутствие внешнего принуждения со стороны других агентов или непреодолимых обстоятельств. Действия, которые агент совершил под угрозой, под физическим давлением или вследствие манипуляции других агентов, не являются свободными. Интуитивно кажется, что понятие свободы предполагает отсутствие принципиальных ограничений, препятствий, сдерживающих факторов со стороны других агентов. Так, свобода слова предполагает отсутствие цензуры. Свобода движения – отсутствие барьеров для движения. Свобода мысли – отсутствие стереотипов, догм и манипулирования в интеллектуальной деятельности. Это – негативная часть определения. Чтобы понятие «ограничение» было осмысленным, нужно определить то, что могло бы быть ограничено. Видимо, свобода воли как минимум предполагает отсутствие принципиальных ограничений реализации желаний агента в поведении. Но все ли действия, реализующие желания самого агента, совершенные без принуждения или ограничения со стороны внешних агентов, свободные? Д. Юм считал, что это именно так. В его понимании, свобода заключается в реализации желаний. В «Исследовании о человеческом познании» он писал: «...мы можем подразумевать под свободой толь-

ко способность действовать или не действовать сообразно решениям воли; если мы хотим оставаться в покое, мы можем это сделать, а если хотим двигаться, то можем сделать и это» [Юм 1996b, 81]. Однако большинство современных аналитических философов не согласны с этим утверждением.

Современные теоретики свободы воли считают, что, даже когда агент беспрепятственно реализует свои желания, свободы воли у него все равно может не быть. Это подтверждает Л. Экстром: «Большая часть философов, однако, считают, что некто может действовать в соответствии с собственными желаниями (а также переживать ментальные состояния желаний), при этом не совершая свободные действия» [Ekstrom 2010, 102]. У этого утверждения несколько оснований. Во-первых, можно сопоставить возможности человека и животных. Животные иногда могут реализовывать свои желания в поведении. Но суждение о свободе воли у животных звучит противоречиво. Второе очевидное основание – различие свободных и несвободных действий у людей. Его можно выявить с помощью лингвистического анализа выражений (1) «Он мог сделать это» и (2) «Он мог бы сделать это, если бы пожелал», как это делает П. ван Инваген. Философ показывает неэквивалентность этих выражений, то есть он доказывает, что нельзя интерпретировать первое высказывание с помощью второго. Для демонстрации этого ван Инваген приводит контрпример. В «Эссе о свободе воли» [van Inwagen

1983, 115] он сопоставляет два утверждения:

А) Смит мог съесть конфету красного цвета;

Б) Смит мог съесть конфету красного цвета, если бы пожелал.

С точки зрения ван Инвагена, различие между этими утверждениями обнаруживается, если допустить, что Смит патологически боится вида крови, а конфеты кроваво-красного цвета. Из совокупности этих обстоятельств следует, что Смит не смог бы съесть красную конфету, хотя смог бы, если бы пожелал. То есть при истинности «Б» «А» оказывается ложным. Значит, нельзя обходиться классическим анализом свободы как возможности поступать сообразно желаниям. Тем более что научные открытия расширяют спектр известных физических причин желаний и действий. Следующий пример, приведенный М. Маккенной и Д. Перебумом в работе «Свобода воли: введение в современную дискуссию» [McKenna, Pereboom 2016, 61], является аналогичным случаем:

Выбор щенка. Отец хочет подарить дочери щенка. В день рождения он приносит ей на выбор белого и черного. Отец решил, что подарит ей того, которого она сама предпочтет. Дочка выбирает черного. Является ли ее действие свободным? Описание действия в этой вымышленной ситуации полностью соответствует требованиям рабочего определения свободы действия. У девочки есть желание, и оно беспрепятственно реализуется в действии. Но что, если девочка

не могла поступить иначе? Допустим, что в вымышленной ситуации девочка боится собак белого цвета. Девочка не может сформировать желание взять второго щенка. Ее желания и, следовательно, действия ограничены. Действие соответствует желанию, но у нее не может сформироваться другого желания. Тогда это не свободное действие? А если девочка в силу своих психических особенностей не могла устоять перед желанием взять первого увиденного ею щенка? Вот она увидела черного и взяла его на руки, но могла увидеть белого, и тогда бы он остался у нее. Выбор черного щенка соответствовал ее желаниям и не был никем навязан. Но был ли это свободный выбор?

Суждения о наличии свободы воли в этом вымышленном сюжете могут показаться спорными. Вдобавок история производит впечатление надуманности. Возможны ли вообще такие специфические психологические особенности? А если невозможны, какую пользу эти примеры несут для прояснения понятия свободы? Поэтому ситуация с выбором щенка только ставит вопрос. Но существует способ прояснить интуиции в отношении свободы и описанного случая. Большинство современных аналитических философов считают, что это можно сделать с помощью понятия ответственности. Ответственность является наилучшим индикатором наличия свободы. Применение понятия ответственности фиксирует случаи, когда мы считаем, что свобода воли существует.

Когда современные аналитики говорят об ответственно-

сти, в частности о моральной ответственности, они «обычно имеют в виду отношение, которое может быть установлено между людьми и совершаемыми ими действиями или между людьми и последствиями их действий» [Talbert 2016, 5]²³. Ответственность связывает людей и их поступки таким образом, что люди, совершившие поступки, становятся объектами особых ответных реакций (*reactive attitudes*)²⁴ со стороны других людей. Этими реакциями в том числе служат негодование, осуждение, похвала, благодарность. Таким образом, под ответственностью можно понимать отношение между агентом и его действием (или последствиями этого действия), которое оправданно вызывает ответные реакции.

Ответственность чаще всего предполагает наличие особого контроля у действующего агента над собственным поведением. Вероятней всего, контроль совпадает с наличием свободы воли. Именно поэтому в современной аналитической

²³ Российские философы часто дают другое определение понятию ответственности, связанное с моральными ценностями. В частности, в книге «Этика» ответственность определяется как «понятие, которое отражает процесс перехода от моральных ценностей и принципов к возникающей на их основе деятельности» [Этика 2016, 157]. Однако для целей настоящего исследования я буду использовать то определение, которое чаще используется в современной аналитической философии.

²⁴ В качестве перевода термина «*reactive attitudes*» я использую русскоязычное словосочетание «ответные реакции». Может показаться, что это тавтология. Однако мне кажется, что здесь это выражение уместно, так как реакция может быть на любое событие, а ответная реакция может быть только на намеренное действие. Именно в качестве реакции на намеренное действие могут быть оправданы такие реакции, как негодование, осуждение или похвала.

философии большая часть теорий свободы воли также является и теориями ответственности. «Моральная ответственность предполагает, что у людей есть некоторый контроль над их действиями. Это связано с тем, что действия, за которые агент несет моральную ответственность, должны быть атрибутируемы ему таким образом, чтобы оправдывать в отношении агента моральные оценки и моральные реакции, такие как осуждение и поощрение. Но как мы должны характеризовать контроль, необходимый для моральной ответственности? Многие философы считают, что такого рода контроль предполагает обладание свободой воли» [Talbert 2016, 14].

О наличии свободы можно судить по наличию ответственности. Поэтому, чтобы определить, в каких случаях наступает ответственность, то есть в каких случаях считается, что человек обладает свободой воли, удобней всего сначала обратиться к судебной практике. В этом отношении судебные прецеденты представляют хороший иллюстративный материал. Они показывают, в каких случаях принято считать человека ответственным, а в каких – нет. С этим согласна и нейрофилософ Патрисия Чёрчленд. По ее мнению, история юридической практики очень полезна для философского исследования, так как выводит дискуссию за рамки «простого академического волейбола» [Churchland 2013, 194] в реальный мир.

В. Действия без принуждения не обязательно являются свободными и ответственными

Современная система правосудия в большинстве стран ограничивает круг тех, к кому может быть применена ответственность, в частности уголовная. Минимальными требованиями к субъекту являются способности понимать, рассуждать и управлять своим поведением. Признать, что субъект несет ответственность, значит также признавать, что субъект может понимать требования закона и морали, обдумывать и принимать решения относительно этих требований и своих действий, а также согласовывать свое поведение с нормами и собственными решениями [Hart 2008, 227]. При определении наличия вины суды учитывают *внешние* ограничения и обстоятельства действий. Одним из освобождающих от ответственности обстоятельств является внешнее принуждение. Однако есть и множество других. Такими обстоятельствами, в частности, служат обстоятельства формирования личности. Так, особенности биографии и воспитания могут смягчать или увеличивать вину. Например, насилие, унижения или психологические травмы, перенесенные в детстве, могут сказаться на характере человека и сформировать у него склонность к антисоциальному поведению, в наличии которой сам он не будет повинен. Поэтому ответ-

ственность в данном случае может быть смягчена.

Суды учитывают и внутренние факторы, ограничивающие контроль и, соответственно, ответственность. Такими факторами являются физиологические и психологические недостатки. Практика учета подобных обстоятельств существует еще со времен Аристотеля [Sapolsky 2004]. Но она значительно расширилась в современной системе правосудия. В американском праве показательна с этой точки зрения защита с помощью ссылки на «невменяемость» подсудимого. Кодификация этой защиты была связана с «правилами Макнотена», введенными в 1840-х. Эти правила устанавливали освобождение от ответственности в случае интеллектуальных недостатков, «дефектов разума» преступника, ведущих к его неспособности понимать природу своих действий или осознавать, что он поступает неправильно.

Правила были введены в связи с убийством Дэниелом Макнотеном Эдварда Драммонда, которого Макнотен принял за британского премьер-министра Роберта Пила. В результате суда Макнотен был освобожден от ответственности. Его признали невменяемым, так как он находился в заблуждении, предполагая, что его преследуют, и не мог препятствовать своим побуждениям совершить убийство, чтобы это преследование прекратить.

Позже, в первой трети XX в., суд округа Колумбия дополнил интеллектуальные критерии невменяемости критериями отсутствия достаточного волевого контроля. От от-

ветственности стали освобождаться люди, действующие по «непреодолимому импульсу». Согласно этому дополнению подсудимый стал признаваться невиновным, если вследствие психического расстройства он не мог справиться с собственным желанием, даже если при этом обвиняемый понимал суть действия и его противоправность, то есть не обладал соответствующими «дефектами разума». Психическая ограниченность стала определяться не только дефицитом рациональности, но и отсутствием волевого контроля²⁵.

Такие принципы невиновности закреплены в правиле Дарема 1954 г. и в Примерном уголовном кодексе 1962 г. Наиболее шумным процессом с применением защиты подобного рода является дело Дж. Хинкли – покушение на президента Р. Рейгана. В ходе спланированного преступления Хинкли произвел шесть выстрелов из пистолета: он ранил президента и трех его сопровождающих. При этом суд признал Хинкли невиновным. В отличие от Макнотена обвиняемый не испытывал заблуждений в отношении Рейгана и его окружения и не считал себя в опасности. Но адвокатам удалось показать, что подсудимый не мог противостоять стремлению совершить преступление. Будучи одержимым страстью к актрисе Джоди Фостер, Хинкли считал преступление единственным способом привлечь ее внимание.

²⁵ По УК РФ – так же: юридическими критериями невменяемости являются два ограничения, интеллектуальное и волевое, причем достаточным для признания лица невменяемым является наличие хотя бы одного из них [Уголовное право России 2004, 118–119].

Впрочем, этот конкретный случай вызывает противоречивые оценки (был ли Хинкли действительно неспособен совладать с собой? не было ли это просто удачной уловкой адвоката?). Вполне возможно, он демонстрирует даже некоторый переко́с в снисходительности правовой системы. Реакция большинства граждан США, наблюдавших за процессом, была негативной. Поэтому он стал причиной отката судебной практики в сторону более жестких критериев невменяемости²⁶. Но вероятно, что откат этот временный. И суды в качестве обстоятельств дела будут принимать доказательства расстройств воли подсудимых, особенно в связи с тем, что в последние несколько десятилетий такие свидетельства становятся более доступными и объективными.

Это следствие прогресса в техническом и концептуальном плане: во-первых, появления технологий неинвазивного сканирования и мониторинга работы мозга (ЭЭГ и ПЭТ), а во-вторых, возможности ассоциировать множество ментальных состояний с нейрофизиологическими. Прогресс предоставил возможность «подглядывать» в субъективный мир агента, давать оценку его когнитивным состояниям и способностям. Стало возможным более объективно судить, мог ли обвиняемый в данном случае сдержаться, поступить иначе. Появился способ изучать непосредственно ключевые обсто-

²⁶ В 1984 г. американский конгресс устранил «дефекты волевого контроля» из критериев невменяемости на уровне федеральных судов в рамках Акта по реформе защиты по невменяемости.

ательства преступления – психологические особенности и состояния мозга обвиняемого.

Исследования показали, какие именно части мозга отвечают за контроль желаний у человека. Оказалось, что контролирующие и сдерживающие функции сосредоточены преимущественно в префронтальной коре (ПФК) головного мозга. Это – место расположения своего рода суперэго, центр цензуры и управления побуждениями. В ПФК происходит выбор «трудных» сценариев поведения: в случаях, когда нужно пересилить автоматический навык или отсрочить получение вознаграждения, выполнить действие более трудоемким, но более эффективным в долгосрочной перспективе способом. В частности, это подтверждается следующими экспериментами. Испытуемому представляется для запоминания список из 16 предметов. Они могут запомнить их двумя способами: самый простой – это попытаться запомнить предметы в представленной последовательности, как есть. Более сложный, требующий усилий, но более эффективный способ – сгруппировать предметы в семантические категории. Эксперимент показывает, что переход ко второму способу всегда требует активации ПФК.

Другой пример показывает роль ПФК в управлении эмоциями. Испытуемым предлагают для просмотра видеоролик с крайне неприятными сценами. В одном случае они должны просто посмотреть этот ролик, а в другом – подавлять естественные отрицательные реакции. Как показывает мони-

торинг, в первом случае ПФК почти не задействована, а во втором принимает активное участие.

Случаи снижения волевого контроля известны давно, хотя связь этих расстройств с дефектами ПФК и лобных долей в целом до развития средств сканирования активности мозга оставалась малоизученной. Самый известный случай подобных нарушений произошел более 150 лет назад с американским железнодорожником Финеасом Гейджем. По свидетельству прессы того времени, при взрывных работах по прокладке путей Гейдж получил тяжелую травму. Металлический прут диаметром 3 см вошел в череп Гейджа ниже левой глазницы и вышел на границе лобной и теменной костей. К значительному удивлению и радости товарищей, Гейдж быстро реабилитировался и уже через два месяца смог вести активный образ жизни. Но его личность сильно изменилась, так что знакомые даже перестали его узнавать. До происшествия Гейджа знали как человека ответственного, уравновешенного, трудолюбивого, справедливого в отношении к подчиненным. После травмы он стал импульсивным, неспособным сдерживать «животные склонности», лишенным чувства уважения к своим коллегам. Он изменился настолько, что, несмотря на физическую реабилитацию, не смог вернуться на постоянную работу. Вероятней всего, причиной такого радикального изменения характера было нарушение связи между левой лобной долей и остальными частями мозга.

Случай с Гейджем вряд ли просвещает нейробиологов. Спустя 150 лет осталось мало свидетельских показаний о ментальных изменениях Гейджа, да и точность сохранившихся сомнительна. Но прецедент сыграл важную роль в концептуальном осмыслении. Этот и подобные случаи показали, насколько нарушения мозга влияют на поведение, на способность контроля. С того времени доказательная база о связи ПФК и контроля эмоциональных импульсов принципиально увеличилась. Подтверждено, что здоровые дети, мозг которых еще находится в состоянии развития (вопреки расхожему мифу, мозг продолжает развиваться даже после подросткового возраста), демонстрируют ограниченные способности контролировать эмоции и импульсивное поведение и выносить моральные оценки. Доказано, что алкоголь угнетает контролирующую функцию ПФК.

Накоплено много прецедентов нарушений части ПФК в результате инсультов. Очень показательное поведение, которое демонстрируют такие пациенты при тестировании: они, с одной стороны, полностью понимают, как рационально следует решать ту или иную задачу, но при этом неспособны справиться с эмоцией, вынуждающей их вести себя нерационально. В ходе одного из нейробиологических тестов пациентам с нарушениями в части ПФК предлагают на выбор два варианта: один, который дает мгновенное, но небольшое вознаграждение, и другой – вознаграждение значительно большее, но с отсрочкой. В ходе тестирования пациенты

обычно комментируют: «Я знаю, я знаю, как работает этот тест. Я должен выбрать этот более сложный путь потому, что вознаграждение в результате будет более значительным. Поэтому я именно это сейчас и выберу...» – и в тот же момент выбирают более «простой» путь [Sapolsky 2004, 1793].

Все это еще раз показывает, что функция волевого контроля может быть расстроена даже при сохранении когнитивных способностей.

И люди с такими особенностями не могут нести ответственность за свои поступки. Это доказывает, что действия без внешних ограничений, соответствующие желаниям агента, не всегда являются свободными и ответственными.

С. Два принципа свободы воли

Что же тогда за компонент нужно добавить к отсутствию принуждения, чтобы получить свободу, необходимую для ответственности? Выше была прослежена связь идеи свободы с отсутствием внешних и внутренних ограничений. Но это определение можно перевернуть с негативного на позитивное и таким образом сделать его более содержательным. Отсутствие ограничений – это не что иное, как наличие возможностей. Тогда получается, что свобода существует, когда есть возможности, когда доступны альтернативы. Вероятно, и случай с выбором одного из двух щенков, и случаи с психологическими дефектами из юридической практики не

подходят под определение свободных действий, так как не предполагают альтернатив для агентов. Следовательно, (1) чтобы агент был свободен в действии и нес моральную ответственность, он должен иметь возможность поступить иначе²⁷. Это утверждение называется *Принципом альтернативных возможностей*, и оно фигурирует во множестве современных работ. Согласно этому принципу, будущее по крайней мере в некоторые моменты представляет собой разветвляющиеся пути или, используя поэтическое выражение Борхеса, сад расходящихся тропок. Делая выбор, агент актуализирует одну из открытых возможностей, вступает на одну из множества тропок. И этот выбор должен быть ему иногда доступен.

Важен, впрочем, не только доступ к выбору, но и связь агента с этим выбором. Выбор должен принадлежать агенту. Эту идею можно выразить в *Принципе автономии* (или «самоопределения»). Этот принцип представляет еще одно важное условие свободы: (2) агент является свободным и, соответственно, морально ответственным, только если причиной его действий «в конечном счете» является он сам. Агент должен обладать контролем над собственными действиями, сам осуществлять выбор, понимая суть альтернатив, не будучи принужден к нему внутренними или внешними обстоя-

²⁷ Аналогичным образом Принцип альтернативных возможностей формулирует Г. Франкфурт в статье «Альтернативные возможности и моральная ответственность» [Frankfurt 2013, 139].

ательствами. Свободное действие должно иметь в качестве причины ментальное состояние, и это ментальное состояние должно принадлежать именно осуществляющему действие агенту. К примеру, свободный выбор во время голосования предполагает не только наличие нескольких кандидатов, но и то, что выбор одного из них исходит от самого агента. При этом агент должен сам понимать, что он голосует, за кого и почему. Так, выбор кандидата, сделанный на основании выпавших в лотерее результатов, не является свободным, как и выбор, продиктованный другим лицом и совершенный под принуждением.

Эти два принципа следует считать наиболее распространенными *условиями свободы воли и ответственности*. Они были обнаружены философами несколько веков назад, но продолжают сохранять актуальность и для аналитиков конца XX – начала XXI в. Эти принципы «представляют собой интерпретацию двух различных аспектов свободы, которые должны быть отражены в любой достоверной теории – а именно принципы самоопределения (или автономии) и доступности альтернативных возможностей. Любая концепция свободного агента должна предполагать возможности и автономию в каком-то смысле...» [Watson 1987, 145], – утверждает Г. Уотсон в знаменитой статье «Свободное действие и свобода воли» (1987). С ним согласна и Л. Экстрем: «Современным теоретикам свободы воли не мешало бы обратиться к весьма полезной статье Г. Уотсона 1987 г. Она придает

структуру дискуссии... и смысл широкому спектру публикаций о проблеме свободы воли» [Ekstrom 2011, 366].

Но, быть может, эти принципы логически связаны, и тогда возможно сократить число требований? Быть может, если агент автономен и сам определяет свое поведение, в то же самое время он имеет альтернативы? Иначе как можно будет определить наличие или отсутствие контроля? Вот агент поступил так, но по собственной инициативе мог поступить иначе. Значит, есть контроль, и агент – источник действия. А если в любом случае он поступил бы так, а не иначе, значит, от него ничего не зависело. Есть, однако, основания считать, что эти принципы несводимы один к другому. Такими основаниями служат контрпримеры – ситуации, когда агент совершает самостоятельный выбор, а альтернатив поведения у него нет. Приведу одну такую ситуацию в пример.

Джон находится дома и занимается собственным делом. Вдруг из соседней квартиры раздается крик и призыв о помощи. Джон слышит этот призыв, однако решает ничего не предпринимать. Ему кажется, что жизнь соседа не так важна, как то, чем он сейчас занимается. Однако предпринять он в любом случае ничего не мог: преступники, напавшие на соседа, предварительно заперли Джона в его квартире снаружи и перерезали все телефонные линии. Таким образом, представляется, что у Джона действительно не было альтернатив, и тем не менее он сам принял решение бездействовать.

Думаю, этот и подобные примеры могут показать, что удо-

влетворение второго принципа возможно без удовлетворения первого, чем демонстрируется, что эти принципы логически независимы. Следовательно, сократить критерии свободы воли до одного условия пока не получится.

1. Свобода воли и детерминизм

Итак, пока предположительно свобода воли заключается в отсутствии принуждения и требует удовлетворения перечисленных выше двух условий (или по крайней мере одного из них). Но могут ли эти условия в принципе быть удовлетворены? Очевидно, что существуют действия, совершаемые под принуждением, действия, детерминированные внешними и внутренними факторами (другими агентами или обстоятельствами), существуют ограничения воли, снимающие ответственность. Но при внимательном анализе становится не совсем очевидно, что воля вообще может быть свободной. Основным препятствием для свободы воли и реализации двух обозначенных принципов является *детерминизм*.

А. Каузальный детерминизм

Детерминизмом называют представление о том, что факты прошлого в сочетании с законами природы определяют наличие одного реально возможного будущего²⁸. Иными сло-

²⁸ По определению ван Инвагена, детерминизм – «это тезис о том, что в любой

вами, это утверждение о том, что причины с необходимостью определяют один результат. Эти две формулировки взаимно дополняют друг друга, так как законы природы и являются по определению тем самым механизмом, который обеспечивает необходимость перехода от событий прошлого к событиям будущего. Это каузальный детерминизм. Есть и другие варианты детерминизма: теологический, генетический, логический. Согласно теологическому детерминизму, все события предопределены замыслом и волей монотеистического божества. Вопросами соотношения теологического детерминизма и свободы воли были заняты средневековые мыслители и схоласты. В генетическом детерминизме утверждается, что поведение всех живых существ определено их генетикой. А тезис логического детерминизма – в том, что все осмысленные предложения языка либо истинные, либо ложные. В том числе предложения о будущем, настоящем и прошлом.

Любой из вариантов детерминизма может представлять проблему для свободы воли. Так, если все события предопределены волей божественного существа, человеку не остается пространства для свободы. Аналогичным образом, пространства для свободы нет, если все реакции организма определены генами. Возможность свободы проблематична

момент существует только одно физически возможное будущее» [Van Inwagen 1983, 3]. Наиболее подробным исследованием понятия «детерминизм» и его соотношения с современной наукой является книга Дж. Эрмана «Учебник детерминизма» [Earman 1986].

также, если все предложения, в особенности о будущем, имеют истинное или ложное значение. Однако эти варианты детерминизма не являются общепринятыми, в отличие от каузального или номологического детерминизма. Каузальный детерминизм (КД) до настоящего времени является распространенной метафизической позицией [Kane 2011a, 6]. И он, по всей видимости, представляет основное препятствие для свободной воли. Это можно показать с помощью двух аргументов, каждый из которых оспаривает соответствующий аспект свободы. Из этих рассуждений следует, что свобода воли не сочетается с необходимостью и детерминизмом.

И эта несовместимость представляет ядро современной проблемы свободы воли. Несовместимость обнаруживается в двух рассуждениях, которые играют существенную роль в современной дискуссии о свободе воли.

В. Два аргумента против свободы в условиях детерминизма

Первое рассуждение основывается на Принципе альтернативных возможностей. Согласно этому принципу, свободным действием является такое действие, которое имеет альтернативы. Допустим, детерминизм верен. Тогда, согласно Тезису о детерминизме, существует только одно реально возможное будущее. События прошлого в сочетании с законами природы с необходимостью определяют будущее, только

одно реально возможное будущее. То есть при данном прошлом дано только одно будущее. Но тогда у действий и желаний агентов нет реальных альтернатив. И это значит, что никакие действия и никакие желания не являются свободными в условиях детерминизма. А если нет свободы, не может быть моральной ответственности.

Впрочем, существует еще один доступный ход – контрфактическая интерпретация «возможности». Допустим, при заданном прошлом будущее определено. Но, если бы прошлое было другим, при тех же законах природы будущее тоже было бы другим, – разве это не альтернатива? Вот в момент t агент хотел произвести действие X и сделал X . Но у действия была альтернатива – действие Y , так как, если бы в момент t агент хотел сделать Y , он бы и сделал Y . Может, такой контрфактический анализ спасает свободу действия? Подобный способ защиты свободы в детерминистическом мире был предложен американским богословом Дж. Эдвардсом, а затем многократно использовался другими философами, в том числе Дж. Муром. Он выражается в интерпретации ситуации «он мог поступить иначе» как ситуации «если бы он хотел поступить иначе, тогда бы он мог поступить иначе». И эта вторая ситуация не противоречит необходимости причин и законов природы.

Однако это, возможно, не выход, так как подобная интерпретация ситуации «он мог поступить иначе» имеет серьезный недостаток. Другая ситуация действительно могла бы

привести к другим действиям. Но другой ситуации быть не могло. Сама ситуация в t была предопределена ситуацией в $t - 1$ и законами природы, и так – до появления данного агента и вообще до появления всех агентов. Наш агент просто не мог иметь другие желания и убеждения и, следовательно, в момент t не мог поступить иначе. Девочка, которой отец предложил выбрать щенка, выбрала черного. Если бы она захотела, она выбрала бы белого. Но в описанном случае она не могла захотеть выбрать белого, и потому, похоже, выбрать белого она не могла. Значение выражения «он мог поступить иначе», скорее всего, подразумевает не только то, что, «если бы он хотел поступить иначе, он мог бы поступить иначе», но и то, что «он мог захотеть поступить иначе»²⁹. А это плохо совмещается с детерминистическим миром. Не больше оптимизма внушает и другое рассуждение. Оно построено на втором принципе.

Свобода, в соответствии со вторым принципом, предполагает контроль или то, что агент сам является автономной причиной своих действий. Но разве возможно такое? Допустим, человек принимает решение что-то сделать. Это его произвольное действие, действие, которое вызвано его желаниями и намерениями. Это действие в качестве причины имеет его предшествующее ментальное состояние. Но это состояние, в свою очередь, имеет другую причину – ментальное или физическое состояние. Регресс причин можно про-

²⁹ Подобный аргумент приводит Родерик Чизом [Chisholm 1964, 6].

должить в прошлое до его детства, младенчества и даже до рождения. Но и здесь цепочка причин не прерывается, ведь каждое событие, в свою очередь, имеет другое событие в качестве причины. Все события в качестве причин имеют события прошлого, в том числе события до появления агента. События прошлого в сочетании с законами природы, согласно детерминизму, определяют все последующие события, в том числе действия агента. Следовательно, все действия в качестве причин имеют события за пределами существования агента. И никакие действия не имеют в качестве *исходной* причины агента. Соответственно, воля человека не является свободной. Причиной конкретного действия является агент и его состояние, но всегда есть и более ранняя причина, которая с необходимостью определяет исход и т. д. Какое тогда значение имеет сам агент? Где здесь пространство для свободы?

Оба аргумента показывают невозможность свободы воли в детерминистическом мире, или несовместимость свободы воли с детерминизмом событий и необходимостью. И эти рассуждения составляют центральную проблему в современных дискуссиях о свободе воли. Теперь мы отчетливо видим вторую опасность на пути к свободе и ответственности. Впрочем, может, проблема решается? Может, она существует только в философской традиции и именно потому, что метафизические установки не прошли должной ревизии?

С. Каузальный детерминизм и квантовая физика

В самом деле, если свобода воли, необходимое условие ответственности, не может сочетаться с детерминизмом, возможно, нет не свободы воли, а детерминизма? Есть некоторые основания так полагать. В то время как философские рассуждения, построенные на эмпирических данных специальных наук (анатомии, нейрофизиологии, психологии и т. п.), склоняют нас к убеждению в детерминизме, современная физика, наиболее фундаментальная из наук, кажется, больше тяготеет к индетерминизму.

Детерминистическое представление о мире стало преобладать в результате успехов ньютоновской механики. Особенно эти успехи были заметны в астрономии. Ньютоновская теория позволила впервые с высокой точностью предсказывать передвижение небесных тел. Успехи в предсказаниях событий, столь отдаленных в пространстве и времени, вселили уверенность, что вскоре можно будет предвидеть любые события и даже состояния всей Вселенной в сколь угодно отдаленном будущем. Мир стал представляться прогнозируемым механизмом, а наука – его новым оракулом. Казалось, что ученым стоит только определить, измерить исходное положение вещей и выявить постоянные законы, управляющие изменениями, и тогда будущее станет откры-

той книгой. Будущее галактик и каждого атома можно будет просчитать, увидеть со стопроцентной определенностью. Если, конечно, позволят вычислительные мощности.

Тем не менее надеждам не суждено было воплотиться. И препятствием стали отнюдь не вычислительные мощности. Физику XX в. ждал концептуальный переворот – открытие квантового мира. Эксперименты, направленные на изучение поведения элементарных частиц, подорвали оптимизм в отношении однозначных предсказаний и показали, что прогнозирование событий гораздо более проблематично. Согласно доминирующей, *копенгагенской интерпретации*, эти проблемы оказались не эпистемологического, а онтологического характера. Ограничения измерений и предсказаний носят не условный, а абсолютный характер. Согласно *принципу неопределенности Гейзенберга*, установить одновременно и с абсолютной точностью позицию и импульс элементарной частицы принципиально невозможно. Следовательно, невозможно однозначно предсказать ее поведение, невозможно определить и траекторию ее движения.

Классическая иллюстрация индетерминизма в квантовых явлениях – эксперимент с двумя щелями. Источник света направляется на светочувствительный экран. Перед экраном установлен щит с двумя вертикальными щелями так, что свет может попадать на светочувствительный экран, только пройдя через эти щели. Кванты света, фотоны, должны проходить через щели и достигать экрана в виде дискретных то-

чечных порций энергии. Для начала закроем одну щель и включим источник света. На экране мы увидим след в виде яркой полосы прямо напротив щели. Результат вполне объясним: часть излучаемых фотонов попадает в щит и не достигает экрана. Фотоны, проходящие через щель, оставляют на экране след в виде проекции щели. Теперь откроем вторую щель. Каков ожидаемый результат?

Можно предположить, что в результате бомбардировки экрана фотонами через щит с двумя щелями получатся проекции двух щелей, две яркие полосы на экране. Но при проверке выходит не так. На экране появляется так называемая *интерференционная картина*, то есть регулярное чередование областей повышенной и пониженной яркости. На экране остаются не две, а множество полос. Но удивительному факту тоже находится объяснение. Если представить, что через щели проходят не корпускулы, частицы материи, а волны, эффект ожидаемый. Ведь волны при наложении одна на другую в зависимости от фазы либо гасят, либо усиливают друг друга. Таким образом, там, где гребни волн совпадают, на экране возникает яркая полоса, а где гребень приходится на ложбину – засветки совсем нет. Такая интерпретация предполагает, что свет проходит одновременно через две щели (образуются как бы две волны). Эта интерпретация заставляет склоняться от корпускулярной природы света к волновой. Но в целом объяснение пока звучит вполне логично. Правда, эта интерпретация успешна только отчасти.

Еще немного усложним эксперимент. Начнем бомбардировать экран не потоком частиц, а поочередно, по одному фотону. В случае поочередной бомбардировки вроде бы нечему создавать интерференцию. То есть мы ожидаем две параллельные засвеченные линии. Но нас ждет еще более удивительный результат. Следы на экране действительно выглядят как результат воздействия частиц. Но и в этом случае в эксперименте получается интерференционная картина, то есть множество световых полос, сочетающихся с абсолютно незасвеченными зонами. А это как объяснить? Что, один излученный фотон покидает источник как частица, становится волной, проходит через две щели, интерферирует сам с собой и попадает в экран вновь как частица? Попробуем это экспериментально проверить. Рядом с одной из щелей установим детектор. Если фотон пройдет через щель, мы получим об этом данные. Итак, начинаем поочередную бомбардировку.

Детектор периодически срабатывает, регистрируя прохождение фотонов через щель. Но на экране вдруг вместо интерференционной картины получаются две полосы. Как будто волновое поведение света вдруг прекратилось под воздействием наблюдения. Фотон обнаружил вмешательство и изменил свое поведение? Но какое отношение к поведению фотона может иметь наблюдатель? Чем объяснить такое странное изменение? И что все-таки такое фотон: частица или волна? Все возможные интерпретации, способные

объяснить эти явления, контринтуитивны. Но наиболее распространена такая: поведение фотона вне наблюдения может быть описано как распределение вероятностей событий. С течением времени эти вероятности эволюционируют. И только в момент регистрации наблюдателем одна из них актуализируется. В квантовой физике этот эффект называется редукцией волновой функции. Без наблюдения не только невозможно определить точное положение частицы в определенный момент. Этого положения вообще нет. Только наблюдение и измерение приводят к его фиксации. Возможно предсказать, какой рисунок получится в итоге на экране, но нельзя с определенностью сказать, где окажется каждая конкретная частица. И это не результат недостаточности знаний: так, видимо, устроен мир.

Странные явления в микроскопическом мире, по всей видимости, управляются вероятностными законами. Будущее элементарных частиц лишено определенности. Тогда что это означает? Что детерминизм ложен? Что найдено пространство для свободы, необходимое для моральной ответственности? И все же оптимизм здесь преждевременен. Только лишь ссылка на квантовую механику, скорее всего, не даст решения проблемы свободы воли. По крайней мере, есть несколько поводов так считать³⁰. Во-первых, в отношении самой теории у физиков нет консенсуса. Существует

³⁰ Подобные основания выделяет Р. Кейн во вступительной статье к сборнику о свободе воли [Кане 2011а, 6–8].

несколько интерпретаций, которые в равной мере согласуются с экспериментальными данными. Действительно, самая распространенная, копенгагенская интерпретация является индетерминистической. Но есть и другие. Например, *эвереттовская интерпретация* (интерпретация многих миров) или *теория де Бройля – Бома* (интерпретация со скрытыми параметрами), напротив, являются детерминистическими. Во-вторых, даже если фундаментальная физика и показывает наличие индетерминизма в мире, специальные науки, напротив, все больше и больше подтверждают детерминизм. Биология, генетика, нейрофизиология, психология выявляют все больше регулярностей, которые определяют поведение живых организмов. Эти науки показывают, что реакции и свойства, первоначально кажущиеся спонтанными и необъяснимыми, имеют в качестве своих причин наследственные, биохимические, эволюционные основания. В таких исследованиях работает допущение, что поведение живых организмов определено, законосообразно и, следовательно, может быть алгоритмизировано. То есть в какой-то момент можно будет создать механизм, программу, способную симулировать все аспекты когнитивной деятельности и поведение. И в отношении нее можно будет делать точные предсказания. Как, возникает вопрос, это сочетается с физикой? Что, не только философы, но и ученые не следят за трендами в фундаментальной науке? Или просто не существует единого поля научного знания?

Допустим, описываемые события в квантовом мире не детерминированы, как события в мире макрообъектов. В квантовом мире действуют статистические законы, а в мире макрообъектов – законы динамические. Означает ли это, что эти миры несовместимы друг с другом? Два мира можно согласовать, если одни описания касаются единичных событий, а другие – статистических. Даже если в отношении единичных событий некоторого типа невозможны достоверные предсказания, они могут быть возможны по отношению к множеству таких событий. Скажем, если один раз бросить монетку, то нельзя с точностью определить, выпадет ли орел или решка. Но если монетку бросать много раз, можно будет достоверно предсказать, что в половине случаев будет выпадать орел, а в половине – решка. Поскольку само существование макрообъектов – это в каком-то смысле статистическое событие для микрообъектов, вывод кажется последовательным. Обычно считается, что квантовые неопределенности должны сглаживаться на макроуровне. И не должны вносить неопределенность и индетерминизм в поведение человека. Это – третье основание. Но и это – не последний повод считать индетерминизм в квантовой физике нерелевантным для решения проблемы свободы воли. Есть еще одно, уже четвертое основание считать, что квантовый индетерминизм не решает философскую проблему свободы воли.

Предположим даже, что квантовые эффекты могли бы усиливаться до событий в макромире. И неопределенность

эволюции состояний микрочастиц распространялась бы на изменения в макрообъектах. Как эта неопределенность соотносилась бы со свободой действия, необходимой для моральной ответственности? С одной стороны, у действий появились бы альтернативы. События в прошлом в совокупности с законами природы не определяли бы будущее однозначно. То есть одно из условий свободы действий было бы удовлетворено – Принцип альтернативных возможностей. Но, с другой стороны, второе условие, обозначенное выше, вряд ли бы выполнялось. Ведь причиной, источником действия, кажется, не был бы сам агент. Реализация той или иной альтернативы в квантовом мире не зависит от воли агентов. Или... все-таки субъект играет роль?

В описанном выше эксперименте было показано, что наблюдения влияют на редукцию волновой функции. Может, и здесь агент имеет значение? Положительный ответ возможен, только если под агентом подразумевать нематериальную субстанцию. Но если агент – это всего лишь особым образом организованный макрообъект, состоящий, в свою очередь, из элементарных частиц, места для свободы воли все-таки не остается. Квантовая неопределенность – это совсем не та свобода, на которую стоит рассчитывать агенту. Фактором, определяющим результат, здесь является случай. Действия, основанные на квантовых эффектах, с точки зрения агента представляют собой непредсказуемые случайности, как конвульсии или судороги. Они неконтролируемы и

не являются тем, что требуется для моральной ответственности.

2. Свобода воли и индетерминизм. Аргумент случайности

Оказывается, не только каузальный детерминизм представляет проблему для свободы воли и моральной ответственности, но и индетерминизм. Ведь индетерминизм, похоже, ведет к потере контроля. Это отражает следующий принцип – *Принцип случайности*: если действия являются индетерминированными, то они являются результатом случайности. А случайные действия не могут быть контролируемые, свободными, и за них агент не может нести ответственность.

На этом принципе основаны аргументы, встречающиеся в философской традиции с давних времен. Они называются Аргументами случайности. Классической формулировкой Аргумента случайности считается формулировка Юма. Шотландский философ так выражает эту идею: «Поступки по самой своей природе суть нечто временное и преходящее; если их не вызывает какая-либо *причина*, коренящаяся в характере и склонностях совершившего их лица, они не могут ни способствовать его славе, будучи хороши, ни бесчестить его, будучи дурны. Поступки сами по себе могут заслуживать порицания, противоречить всем правилам нравствен-

ности и религии, но человек не ответствен за них: поскольку они не были вызваны каким-либо постоянным и непреходящим свойством характера и не оставляют такого же рода последствий, человек не может подвергаться за них наказанию или мщению. Таким образом, согласно принципу, отрицающему необходимость, а следовательно, и причинность, человек, совершив самое ужасное преступление, остается столь же чистым и незапятнанным, как и в момент своего рождения...» [Юм 2009, 99]. Итак, свобода воли предполагает устранение необходимости действия. По мнению Юма, необходимость есть неотъемлемая характеристика причинности. Устраняя необходимость, свобода устраняет и причинность, а вместе с ней – и связь с характером человека, делая действие случайным. Соответственно, человек не может и не должен нести ответственность за действия, определенные свободой воли.

В современной аналитической философии также используются аргументы, основывающиеся на Принципе случайности. В частности, они встречаются в работах И. Хаджи [Haji 2001], А. Мили [Mele 2008] и П. ван Инвагена [van Inwagen 2000]. Различные формулировки отличаются в том, как они определяют случайность. Некоторые авторы считают, что «случайность» – это отсутствие возможности объяснить событие. Из этого они делают вывод, что индетерминированные события являются случайными потому, что у них отсутствует соответствующее объяснение. Другие философы

считают, что понятие «случайность» не требует редукции и само по себе очевидно. Эти авторы оставляют понятие «случайность» без расшифровки, а утверждение, что индетерминированные действия являются случайными, принимают как самоочевидное. Однако исходная интуиция в аргументе сохраняется. Кроме того, современные версии аргумента представляют важное усовершенствование.

Один из вариантов современного Аргумента случайности использует понятие возможных миров для интерпретации модальности. Согласно этой интерпретации, если некоторое событие возможно, оно «существует в ближайшем возможном мире», а если событие необходимо, то оно «существует во всех возможных мирах». Аргумент звучит так:

1. Свобода воли заключается в отсутствии детерминации действий предшествующими событиями.

2. Пусть в реальном мире W человек, обладающий свободой воли, в момент t совершает поступок A .

3. И пусть в ближайшем возможном мире W^1 (полностью идентичном W до момента t) этот же человек, обладающий свободой воли, в момент t совершает поступок B .

4. Поскольку A и B – свободные действия, не существует причин, которые с необходимостью предопределяли бы действие A или B .

5. Ничто из тех способностей агента, его ментальных состояний, характеристик, диспозиций или мотивов, что предшествовали t , не является причиной действий

A и *B* и не объясняет выбора действий.

6. Следовательно, осуществление *A* или *B* – дело случая. И человек не отвечает за то, что в *W* он в момент *t* сделал *A*, а не *B* (как и за то, что в *W*¹ он совершил *B*, а не *A*), поскольку он не может нести ответственности за случайные события.

Этот аргумент можно проиллюстрировать гипотетической ситуацией. Представим, что Джон размышляет, где ему провести отпуск: на своей даче или в экспедиции на остров Тристан-да-Кунья. После значительных размышлений, во время которых он взвешивает все за и против, он в итоге выбирает поездку на собственную дачу. Из тезиса индетерминизма следует, что некоторые события происходят без достаточных причин. То есть одни и те же события могут развиваться по-разному. Это удачно иллюстрирует метафора разветвляющихся тропинок. Соответственно, если выбор Джона был индетерминированным, после всех тех же самых размышлений он мог бы выбрать экспедицию. Одни и те же размышления могли привести его к совершенно противоположным действиям, на другие тропинки. Следовательно, Джон не отвечает за свой поступок. И никакие ссылки на размышления Джона не могут объяснить его выбор. Отсутствие детерминизма, исходя из этой ситуации, приводит к отсутствию ответственности.

Аргумент и гипотетическая ситуация в целом соответствуют интуициям Юма. Однако современная версия более

опасна для свободы. Опасность в том, что в ней рассматриваются не только любые индетерминированные поступки, но и поступки намеренные. В мысленном эксперименте Джон всерьез размышляет над выбором, имеет рациональные основания в пользу альтернатив. И при этом он может решить поступить так или иначе. То есть за гипотетической ситуацией стоит аргумент, где посылки 2 и 3 оснащены дополнительными обстоятельствами.

2. Пусть в реальном мире W человек, обладающий свободой воли, в момент t *намеренно и по собственному желанию* совершает поступок A .

3. И пусть в ближайшем возможном мире W^1 (полностью идентичном W до момента t) этот же человек, обладающий свободой воли, в момент t *намеренно и по собственному желанию* совершает поступок B .

Эта модернизация делает аргумент более устойчивым к критике. Однако не будем забегать вперед. Пока рано подробно разбирать специфику и преимущества модернизированного аргумента. Даже классические, юмовские возражения против сочетаемости индетерминизма и ответственности представляют проблему.

3. Дилемма свободы воли: соотношение с детерминизмом и индетерминизмом

Из изложенных выше рассуждений следует, что особенности законов квантовой физики, вероятно, не создают прочного фундамента для свободы и не спасают моральную ответственность от детерминизма. Более того, они приводят к еще одному поразительному выводу: свобода и моральная ответственность, кажется, несовместимы не только с детерминизмом, но и с индетерминизмом. Индетерминизм, похоже, превращает действие в случайное событие, трансформирует ментальное в эпифеноменальное. Индетерминизм предполагает, что причины не определяют действия однозначно. Но в таком случае осуществление некоторого действия или его альтернативы оказывается делом случая. Если действия происходят случайно, то агент не является их источником. И потому он не может нести какую-либо ответственность. Ведь, как было отмечено выше, причинность и контроль – родственные понятия. Итак, обнаруживается дилемма: моральная ответственность предполагает свободу воли. Свободная воля (1) предполагает возможность реальных альтернатив действий. А также что (2) агент является исходной причиной действий. Согласно детерминизму, агент, видимо, не может быть причиной действий в конечном счете, и у действий не может быть реальных альтернатив. Та-

ким образом, детерминизм, похоже, несовместим с моральной ответственностью. Но индетерминизм также несовместим с моральной ответственностью, так как в отсутствие каузальной необходимости агент перестает иметь связь со своим поведением. Он перестает быть исходной причиной поведения. Исчезает контроль как основание ответственности за поступки. Возможно, даже исчезает само действие вместе с ментальной каузальностью. Так понимаемая индетерминистичная свобода трансформирует действие в случайное событие. Это – порочный круг. С одной стороны – необходимость, с другой – случайность. И то и другое является препятствием для возможности свободы и ответственности. Как резюмирует известный российский этик А. Разин, «в истории философии в понимании свободы издревле было выявлено определенное противоречие. Оно связано с вопросом об обусловленности человеческой воли. Если воля совершенно ничем не обусловлена... свобода в таком случае просто случайный выбор наугад. Но если воля сама по себе чем-то обусловлена, она перестает быть свободной» [Разин 2015, 9]. Если бы только агент мог быть *causa sui*, причиной самого себя и причиной своих ментальных состояний, а его ментальные состояния – причиной действий, тогда моральная ответственность, видимо, могла бы существовать. Но это несовместимо с современным представлением о каузальности.

4. Альтернативные позиции по вопросу свободы воли конца XX – начала XXI В

Впрочем, философы не считают проблему нерешаемой. В современной философии представлены практически все альтернативы решений. Есть множество крайне утонченных вариантов распутывания этого концептуального узла. Современные варианты решений проистекают из философских позиций, оформившихся еще в Новое время. Но было бы совершенно неправильно говорить об отсутствии прогресса в этой области. Позиции по проблеме свободы воли были значительно усовершенствованы, особенно с середины XX в. Прогресс был связан в первую очередь с работами П. Стросона, Г. Франкфурта и П. ван Инвагена. И хотя консенсус так и не сложился, современные теории представляют явный шаг вперед в сравнении со своими предшественницами.

Демаркационной линией в философских позициях служит отношение к детерминизму. Часть философов считает мир детерминистичным, а часть отстаивает индетерминизм. Вторая черта, в соответствии с которой разделяются позиции, – отношение между детерминизмом и свободой воли. Пожалуй, это даже более важная черта. Компатибилистами (сторонниками теории совместимости) называют тех, кто считает свободу воли совместимой с детерминизмом; инкомпатибилистами – тех, кто убежден в их несовме-

стимости. Компатибилизм можно также разделять на классический и современный. Классическим компатибилизмом называется позиция, основанная на определении свободного действия как действия, соответствующего желанию. То есть классические компатибилисты считают, что для ответственности достаточно ментальной каузальности. Классический компатибилизм, как было показано, уязвим для контрпримеров, подобных тем, что приводят ван Инваген, Перебум и Мили. Не только препятствия желаниям, но и сами непреодолимые желания могут представлять ограничения для свободы. Поэтому современные компатибилисты (Дж. Фишер, Д. Деннет, Г. Франкфурт) повышают требования к свободному действию. Среди компатибилистов можно выделить также суперкомпатибилистов. Этот новый термин вводит в обращение М. Варгас [Vargas 2012]. Согласно этой позиции, свобода и ответственность совместимы как с детерминизмом, так и с индетерминизмом. Суперкомпатибилизм – наиболее распространенная сегодня позиция среди теоретиков свободы воли. Это подтверждает К. Тимпе: «... современные компатибилисты – это сторонники свободы “в любом случае”; иными словами, они считают, что у нас есть свобода воли, если детерминизм истинен... но также и если детерминизм ложен» [Timpe 2013, 175]. Сторонники суперкомпатибилизма считают проблему Случайности в случае индетерминированных действий решаемой.

Инкомпатибилистские позиции можно разделить на ли-

либертарианство и жесткий инкомпатибилизм. Жесткий инкомпатибилизм – это позиция, согласно которой свободы и ответственности не может быть в случае как детерминизма, так и индетерминизма. Эту позицию можно назвать самой пессимистической в отношении свободы воли. Согласно либертарианской позиции, индетерминизм верен и он обеспечивает возможность свободы и ответственности.

Либертарианство можно разделить на дуалистическое и натуралистическое, или физикалистское. Дуалистическое либертарианство допускает отсутствие каузальной замкнутости физического мира и возможность воздействия со стороны нефизических сущностей. Такие теории менее всего распространены в настоящее время. Зато есть варианты натуралистического либертарианства, вдохновленного вероятностными законами квантовой физики. Выше я попытался показать, что на первый взгляд квантовая физика не дает «пространства для свободы агента». Но есть попытки построить теории свободы воли на базе вероятностных законов и обойти Аргумент случайности. Одной из наиболее успешных и проработанных в этом отношении является теория Роберта Кейна³¹.

³¹ Среди современных аналитических философов либертарианскую позицию также занимают К. Франклин и М. Балагер.

II. Решение проблемы свободы воли в либертарианстве Р. Кейна

1. Этапы либертарианской кампании

Как считает Р. Кейн, перед сторонником либертарианства стоят две основные задачи. Сначала ему нужно показать, что свобода, необходимая для моральной ответственности, несовместима с детерминизмом. (Эта задача, впрочем, общая как для либертарианца, так и для сторонника жесткого инкомпатибилизма.) А потом философ-либертарианец должен продемонстрировать, как свобода и моральная ответственность сочетаются с индетерминизмом. Первая задача носит критический характер. Либертарианцу нужно показать, что существует как минимум один вид свободы (и, соответственно, ответственности), который не может быть реализован в условиях детерминизма. И что именно этот вид свободы важен. Вторая задача – это развитие непротиворечивого представления об этом виде свободе и ответственности в условиях индетерминизма. Как было показано выше, простое отрицание детерминизма еще не создает условий для реализации свободы. Чтобы либертарианская концепция была конструктивной, она должна продемонстрировать, как свободные действия могут направляться и контролиро-

ваться агентом. Защитник либертарианской позиции должен пройти между Сциллой необходимости и Харибдой случайности.

Кейн сравнивает два этапа в развитии либертарианской теории с восхождением и спуском с горы. Недостаточно только подняться на вершину, нужно суметь с нее спуститься. Если решить только одну проблему, штурмом взять вершину инкомпатибилизма, но задержаться там наверху – понятие свободы останется чем-то мистическим, непроясненным. Разреженный воздух способствует иллюзиям и галлюцинациям. В истории философии некоторые уже становились жертвами этой «панической метафизики». Многих философов, как пишет Кейн, посещали «видения ноуменальных Я, неслучающихся причин (*nonoccurrent causes*), трансэмпирических эго и прочие фантазии» [Kane 1996, 14]. Одним из вариантов такой панической метафизики является апелляция к субстанциональной каузальности. Это – довольно экстравагантный вид либертарианства. Он приводит к отрицанию уже известного нам Принципа каузальной замкнутости физического мира. Но это крайне нежелательно. Сам Кейн стремится избежать таких специфических решений. Он рассчитывает не только подняться, но и успешно спуститься с горы, не став жертвой опасных иллюзий. И в этом походе у него есть проводники – это обыденные интуиции и современные физические теории.

Кейн считает, что бытовые интуиции об ответственности

и свободе воли – ближе всего к либертарианской позиции. Поэтому большинство людей, не задумывающихся всерьез над философскими проблемами, – либертарианцы. Размышляя о будущем, люди представляют альтернативы реальными. Совершая выбор, они верят, что могут решить и так, и иначе, то есть представляют будущее как «сад расходящихся тропок». Но в детерминированном мире будущее напоминает скорее многополосное шоссе: множество не пересекающихся между собой прямых линий. То есть интуитивные представления о свободе несовместимы с детерминизмом. Детерминизм также несовместим и с интуитивными представлениями об ответственности.

Интуитивно кажется, что, чтобы быть ответственным, человек сам должен быть причиной своих поступков и создателем собственного характера. Как только мы узнаем, что поведение человека было полностью обусловлено некоторыми сторонними причинами, мы снимаем с него ответственность. То есть обыденная точка зрения предполагает, что в случае свободных действий цепочка причинности должна где-то разрываться, что в детерминации должны быть лакуны и что агент способен как бы левитировать в каузальном мире. Поэтому Кейн может опираться на интуицию и бытовые представления. Но это – не единственная его опора.

Вторым важным союзником для Кейна является современная фундаментальная физика. Выше я уже отмечал, что квантовый мир, по крайней мере, согласно некоторым ин-

терпретациям, не является детерминированным. То есть события в нем происходят только с определенной вероятностью, но не могут быть предсказаны наверняка. Анализ, однако, показал, что одни только положения квантовой физики не позволяют считать макромир индетерминистичным. К тому же встает вопрос о том, какое отношение агент имеет к квантовой неопределенности. Нельзя ведь отвечать за события, которые совершаются по случайности. И тем не менее фундаментальная наука дает концептуальные инструменты для современного либертарианца. В первую очередь она позволяет говорить о вероятностной каузальности. А вероятностная каузальность совсем не то же, что простая случайность. Стороннику либертарианской позиции остается продумать связь между вероятностной каузальностью и агентом. И тогда решение проблемы совместимости свободы и индетерминизма может быть обнаружено.

2. Свобода воли и детерминизм несовместимы

Полагаться на одни только интуиции или феноменальный опыт в вопросе о свободе воли ненадежно. Как и в случае с ментальной каузальностью, простой демонстрацией или ссылкой на очевидность здесь не обойдешься. Недостаточно просто поднять руку и сказать: «Вот я поднял руку, а мог бы ее не поднимать. Значит, свобода воли существует». Несмот-

ря на привлекательность поэтической метафоры сада расходящихся тропок по сравнению с индустриальным вариантом параллельных шоссе, предпочтение одного или другого варианта для описания процесса выбора не очевидно. Поэтому Кейн опирается на аргументы. Два основных пути аргументации в пользу инкомпатибилизма были обозначены выше: первый – через Принцип альтернативных возможностей, второй – через Принцип автономности, или «Принцип ответственности в конечном счете» (так называет его Кейн). Анализ свободы Кейн начинает с Принципа альтернативных возможностей. На нем построен один из ключевых аргументов современных инкомпатибилистов – Аргумент последствий.

А. Принцип альтернативных возможностей

а) Аргумент последствий

Р. Кейн называет Аргумент последствий «одним из самых обсуждаемых аргументов о несовместимости свободы и детерминизма за последние двадцать пять лет» [Капе 1996, 44]. Истоки этого рассуждения могут быть найдены еще у шотландского теолога Дунса Скота. Но в современном, более точном виде он был сформулирован К. Жине (1966), Д. Виггинсом (1973) и П. ван Инвагеном (1975). Классическая разработка этого аргумента представлена в работе Питера ван Инвагена «Эссе о свободе воли» [van Inwagen 1983].

Обычно рост популярности и активизация дискуссий вокруг Аргумента последствий считается именно его заслугой. Ван Инваген рассуждает: «Если детерминизм верен, тогда наши действия являются *последствиями* законов природы и событий в удаленном прошлом. Но от нас не зависит то, что произошло до нашего рождения; и от нас не зависят законы природы. Поэтому и последствия событий и законов (включая наши текущие действия) не зависят от нас» [van Inwagen 2013, 117]. Для удобства анализа это рассуждение можно разложить на несколько посылок и вывод.

1. Мы не имеем возможности изменить прошлое и законы природы.

2. Если детерминизм верен, наши текущие действия являются необходимыми последствиями событий прошлого и законов природы.

В. Следовательно, мы не имеем возможности сделать так, чтобы наши текущие действия были другими.

На первый взгляд первая посылка кажется достоверной. С прошлым и законами природы действительно нельзя ничего сделать. Вторая посылка опирается на утверждение о каузальном детерминизме, следует из его определения. Опровергать ее не имеет смысла. Из этих посылок логично следует вывод: в детерминистическом мире наши действия являются последствиями не зависящих от нас событий, необходимым результатом действия законов природы и событий прошлого, в том числе событий, произошедших до наше-

го рождения. Значит, мы не можем делать что-то другое, нежели то, что делаем. То есть аргумент исключает наличие альтернативных возможностей. Но если альтернативные возможности отсутствуют, то, согласно Принципу альтернативных возможностей, отсутствует и свобода воли. Так из Аргумента последствий следует отсутствие свободы. Но действительно ли свобода предполагает наличие альтернативных возможностей? Действительно ли истинен этот принцип?

б) Критика Принципа альтернативных возможностей с помощью примеров Франкфурта и Деннета

Кейн признаёт, что сам Принцип альтернативных возможностей в целом и та его форма, в которой он представлен выше, в частности в последнее время нередко оспариваются. Атака на этот принцип началась в 1969 г. в связи с публикацией ключевой статьи Г. Франкфурта «Альтернативные возможности и моральная ответственность» [Frankfurt 2013] и была подхвачена Д. Деннетом в монографии «Пространство для свободы: типы свободы воли, которые стоит желать» [Dennett 1984]. Франкфурт и Деннет поставили под сомнение то, что раньше вызывало общий консенсус, – необходимость альтернативных возможностей как условия свободы. Они представили примеры реальных и гипотетических, но допустимых ситуаций, в которых альтернативные возможности отсутствуют, но при этом кажется, что агент по-

ступает по собственной воле и несет ответственность за свои поступки. Наибольший резонанс получили примеры Франкфурта.

Пример 1. В первом примере Франкфурт предлагает вообразить ситуацию, в которой Джонс совершает поступок А. По замыслу автора, Джонс самостоятельно принимает решение совершить А. Однако существует и другая причина в пользу поступка А – внешняя угроза. После того как Джонс принял решение осуществить А, ему угрожают, что, если он не совершит А, с ним, его семьей и близкими жестоко расправятся [Frankfurt 1988, 2–3]. Фактически у Джонса нет альтернатив – он вынужден делать А. При этом интуитивно кажется, что Джонс несет моральную ответственность за поступок А. Ведь, несмотря на наличие угрозы, его действующим мотивом было собственное желание. И даже если бы внешней угрозы не существовало, он все равно сделал бы А. Франкфурт пишет: «Кажется очевидным, что в этих обстоятельствах тот факт, что Джонсу угрожали, никоим образом не снимает с него моральной ответственности, которую в противном случае он нес бы за свое действие» [Frankfurt 1988, 3]. Можно, однако, возразить, что у Джонса все-таки была альтернатива, ведь так или иначе он мог не подчиниться угрозам и сделать не А, а Б. Что ж, если в этом случае наличие реальной альтернативы все-таки есть, можно далее усложнить ситуацию так, чтобы альтернатива полностью испарилась. Так сконструирован второй пример Франкфурта.

Пример 2. Блэк хочет, чтобы Джонс выполнил действие А. Блэк имеет возможность заставить Джонса выполнить А, но он не хочет вмешиваться раньше времени. Поэтому он ждет, что решит делать сам Джонс. Если Джонс сам выберет А, Блэк не станет ничего предпринимать. Если же Джонс попытается выбрать В, то Блэк вмешается и заставит его выполнить А. Можно представить, что у Блэка для этого есть все необходимые условия. Так, пока Джонс спал, Блэк установил в мозг Джонса радиоуправляемые электроды, которые при необходимости простимулируют последнего сделать А. У Джонса просто нет альтернатив. И при этом, если он сам выбирает А, он, видимо, будет нести ответственность. Он несет ее в той же мере, как если бы Блэка вообще не существовало и у него были бы реальные альтернативы. Джонс может являться морально ответственным агентом, объектом поощрения или порицания, хотя фактически альтернатив у него нет.

Франкфурт апеллирует к гипотетическим ситуациям, а Деннет предлагает обратиться к историческому событию – знаменитому выступлению Лютера, с которого началась протестантская реформация. В ходе выступления перед судом обвиняемый в ереси Лютер подтвердил свои убеждения, выраженные им в сочинениях, словами: «На том стою. И не могу иначе». Лютер утверждал, что не имеет альтернатив, и, вполне вероятно, он выражался буквально. Но разве это была попытка уклониться от моральной ответственности за

свой выбор или жалоба на отсутствие свободы? Разве этим высказыванием он ставил под сомнение наличие контроля? Наоборот, в этих высказываниях он принимал ответственность, выражал полное единство мысли и действия. Его позиция детерминировала его решение, не оставляя альтернатив. Но необходимость и отсутствие других возможностей ни в какой мере не освобождали его от ответственности. Деннет утверждает: «Его декларация – это свидетельство того, что мы просто не освобождаем от порицания или похвалы человека за действие, полагая, что он не мог поступить иначе» [Dennett 1984, 133].

Примеры Франкфурта и Деннета предназначены для фальсификации Принципа альтернативных возможностей – принципа, согласно которому «человек морально ответствен за свои действия только в том случае, если он мог поступить иначе». И эту роль они выполнили достаточно эффективно. Помимо сомнений в Принципе альтернативных возможностей есть и другие основания не принимать Аргумент последствий.

с) Значение выражения «иметь возможность»

Большинство противников Аргумента последствий считают, что проблема этого аргумента – в заключении. По сути дела, они атакуют утверждение о том, что из неизменности детерминистических законов природы и прошлого следует отсутствие альтернативных возможностей. Альтернативные возможности в каком-то смысле сохраняются у нас даже в

детерминистическом мире. Это можно проследить на практическом примере. Я сейчас пишу, а не читаю. Но ведь я мог бы читать. Я мог бы сейчас читать, рассказывать по комнате или пить чай. Все это является моими актуальными возможностями как полноценного агента и, кажется, согласуется как с законами природы, так и с событиями прошлого. А если аргумент приводит к тезису о том, что это невозможно, скорее всего, либо аргумент ложен, либо понятие «иметь возможность» в аргументе задано в каком-то чрезвычайно узком, специализированном значении. Но тогда и область его применения минимальна. Похоже, что Аргумент просто вводит такое значение понятия «иметь возможность», которое несовместимо с детерминизмом. При этом такая трактовка не является единственно возможной. Тем хуже для инкомпатибилиста. Детерминизм может быть несовместим с понятием «иметь возможность», но совместим с понятием «иметь возможность». И как теперь разобраться, какое из понятий необходимо для свободы воли и моральной ответственности? Кажется, впереди патовая ситуация. Одни настаивают на узком определении «возможности», а другие – на широком, опять прибегая к условной интерпретации. Где искать выход?

Р. Кейн считает, что пата в перспективе можно избежать. Но именно в связи с этой патовой ситуацией он предлагает ориентироваться в первую очередь не на Принцип альтернативных возможностей, а *на ответственность в конеч-*

ном счете, то есть на Принцип автономии. «Фокусирование только лишь на “альтернативных возможностях” (или “разветвляющихся тропках” в будущее) или “возможности поступать иначе”, как в случае с Аргументом последствий, – слишком слабое основание, чтобы построить на нем доказательство несовместимости свободы воли и детерминизма» [Kane 2007, 13]. В пользу такого решения, по его мнению, говорят также примеры Франкфурта и Деннета. «Так, я согласен с Франкфуртом, что его Принцип альтернативных возможностей ложен» [Kane 2011b, 384], – пишет он. Согласно Кейну, ошибка этого принципа состоит в том, что он требует, чтобы *каждое* свободное действие имело альтернативы. Но, как показывают контрпримеры, этого не требуется. Франкфурт и Деннет, по мнению Кейна, действительно уловили важный момент. Но Деннет и Франкфурт не правы в том, что вообще отрицают необходимость альтернативных возможностей для свободы воли. Кейн считает, что они все равно нужны, хотя бы для некоторых решений. Эти решения определяют личность и делают ее свободной. Доказательство этого следует из Принципа автономии.

В. Принцип автономии и «ответственность в конечном счете»

Вторым условием свободы, как было установлено выше, является автономия личности. Она выражается в возможно-

сти агента быть исходной причиной действий. Все обычные причины, являющиеся частью каузальных цепочек, не подходят для этого. Их истоки уходят в бесконечное прошлое. Для свободы, вероятно, необходимо, чтобы агент не зависел от причин, уходящих в бесконечное прошлое. Только в таком случае он сам будет устанавливать выбор и действительно контролировать свое поведение. Уже в ранних работах Кейна учитывается это условие свободы. Философ обозначает его как требование «абсолютного суверенитета» агента по отношению к действию (*Sole* или *Ultimate Dominion*). В более поздних работах Кейн называет это условие «ответственностью в конечном счете» (ОКС). По определению автора, согласно Принципу ОКС, «чтобы быть ответственным в конечном счете за действие, агент должен отвечать за все, что является достаточной причиной или мотивом действия» [Kane 2007, 14]. Так, если основанием и объяснением действия агента являются его интенции, агент должен нести ответственность за формирование интенций. Если основанием действия является характер, агент должен нести ответственность и за формирование своего характера.

Принцип ответственности в конечном счете менее уязвим для критики. И его нельзя подорвать примерами Франкфурта или Деннета. Ведь этот принцип не требует наличия альтернативы для каждого ответственного действия. Агент, по мнению Кейна, может отвечать за свои поступки даже тогда, когда перед ним – действительно единственно возмож-

ное действие. Важно только, чтобы он отвечал за все, что является достаточной причиной или мотивом этого действия. В этом контексте Кейн объясняет и ситуацию с высказыванием Лютера: когда Лютер говорил «На том стою. И не могу иначе», он действительно мог не иметь альтернатив, не быть в состоянии отказаться от собственных текстов и идей. Но он не снимал с себя ответственность. Его выбор был обусловлен твердой позицией, которая, в свою очередь, основывалась предыдущим мучительным выбором, сделанным в результате колоссальных сомнений и неопределенности. Высказывание Лютера – это выражение обдуманной и определенной позиции, сформировавшейся как результат трудного, но свободного выбора. И речь здесь не идет только лишь о внутренней психологической борьбе и когнитивной неопределенности, которую ему пришлось пройти. Речь идет о *метафизической неопределенности и сознательном выборе в автобиографическом прошлом Лютера.*

Именно моменты метафизической каузальной неопределенности Кейн считает основанием моральной ответственности личностей в дальнейшей их биографии. Они позволяют, по его мнению, остановить регресс от ближайших причин к причинам более ранним. И в жизни каждой личности такие моменты имеют особое значение. Они являются характерообразующими: «...*свобода воли* – в отличие от только лишь *свободы действия* – заключается в формировании характера и мотивов, которые являются источниками и осно-

ваниями похвалы и порицания добродетельных и порочных деяний. Свобода воли... состоит в самосозидании» [Кане 2007, 16]. Только в эти моменты агенту необходимо наличие доступных альтернатив.

Таким образом, путь Кейна к инкомпатибилизму можно определить так:

1. *Ответственность в конечном счете* требует, чтобы агент отвечал за достаточные причины своих действий.

2. Достаточные причины действий агента могут уходить бесконечно в прошлое, если только не существует действий на жизненном пути агента, которые не имели достаточных причин.

3. Детерминизм – это положение о том, что из фактов прошлого и законов природы с необходимостью следуют все последующие события. (Определение детерминизма)

4. Если существуют действия без достаточных причин, то детерминизм ложен.

Этот путь следует называть *Аргументом от ответственности в конечном счете*. Его исходной посылкой является Принцип ОКС. Этот принцип приводит Кейна к каузальному регрессу: переходу от непосредственных причин действия к предшествующим причинам. Если допустить бесконечный регресс, ответственность в конечном счете невозможна, так как причины действий агентов будут уходить все дальше и дальше в прошлое до момента, когда агента не су-

ществовало. А агент не может отвечать за события, которые произошли до того, как он начал существовать. Следовательно, чтобы ответственность вообще была возможна, регресс должен быть остановлен где-то в интервале жизни агента, где и должна быть найдена исходная причина действия. Но такой причиной могут быть только другие действия агента, у которых не было достаточных причин. Это действия, которые не следуют с необходимостью из событий прошлого и законов природы. Но если подобные действия должны существовать, то детерминизм должен быть ложен³². Таким образом, Кейн, по его мнению, доказывает, что детерминизм несовместим с ответственностью в конечном счете и со свободой воли.

С. Соотношение двух принципов свободы в теории кейна

Из доказательства видно, что Принцип ответственности в конечном счете играет решающую роль в рассуждениях Кейна. И в его аргументе в пользу инкомпатибилизма нет ссылок на альтернативные возможности. Поэтому Кейна причисляют к «инкомпатибилистам по источнику» (*source incompatibilist*). Но это не означает, что философ отрицает Принцип альтернативных возможностей и его необходимость для свободы. Это отличает его от сторонни-

³² Аргумент в имплицитном виде содержится в [Кане 2007, 14–15].

ков так называемого «узкого инкомпатибилизма» (*narrow incompatibilism*). Последние критерием свободы воли считают только один принцип – Принцип ОКС. А Кейн утверждает, что как ОКС, так и Принцип альтернативных возможностей необходимы для реализации свободы воли и основания ответственности. Но при этом, согласно ему, Принцип ответственности в конечном счете – это базовый принцип, а Принцип альтернативных возможностей – производный. Такую их связь Кейн доказывает с помощью набора вымышленных ситуаций, которые называет *примерами в стиле Остина*. Эти ситуации противоположны примерам Франкфурта и Деннета. В них у агента есть альтернативные возможности, но нет свободы и ответственности. Они показывают, что, даже если Принцип альтернативных возможностей необходим для свободы воли, он явно недостаточен.

В первой ситуации предполагается, что для выигрыша чемпионата по гольфу игроку требуется выполнить патт с трех футов – катящий удар с близкого расстояния. Но в связи с нервной судорогой игрок промахивается. Во второй ситуации киллер совершает покушение на премьер-министра, стреляя из снайперской винтовки, но в связи с конвульсивным движением промахивается и убивает помощника министра. В третьей ситуации Остин сам намерен выбрать в кофе-машине приготовление кофе без сливок, но в связи с неожиданной дисфункцией мозга случайно нажимает кнопку для приготовления кофе со сливками. Во всех трех слу-

чаях можно допустить, что в цепочке событий присутствует настоящая случайность. Представим, что в нормальные казуальные процессы вторгается квантовая неопределенность. То есть все описанные события – промах, убийство помощника и приготовление кофе со сливками – недетерминированы, то есть не имели достаточных причин. И эти события, соответственно, могли иметь реальные альтернативы. Игрок мог попасть в лунку, убийца – убить премьер-министра, а Остин, желавший кофе без сливок, мог бы получить кофе без сливок. Принцип альтернативных возможностей удовлетворен.

Но считаем ли мы, что все эти действия были свободными? – спрашивает Кейн. И были ли они добровольными? Хотел ли игрок в гольф промахнуться, убийца – убить помощника министра, а Остин – выпить кофе со сливками? Кажется, все эти действия были произвольными и несвободными. Агенты, совершившие эти действия, не выбирали их, а совершили их по случайности или ошибке. Следовательно, делает вывод философ, индетерминизма в сочетании с альтернативными возможностями недостаточно для свободы воли. Для свободы воли агенты должны иметь возможность самостоятельно управлять своей волей, быть ее источником и исходной причиной. А это условие не выполняется в представленных ситуациях.

И все-таки альтернативные возможности необходимы для свободы. Кейн рассуждает так: свобода воли предполагает

ответственность в конечном счете за волевые акты и действия агентов. В свою очередь, как следовало из предыдущего аргумента, ответственность в конечном счете требует действий, не продиктованных необходимостью. Эти действия являются *волеустанавливающими* (*will-setting actions*) (ВУД). Неопределенность, предшествующая волеустанавливающим действиям, должна быть желаемой. Но если неопределенность должна быть желаемой, значит, отчасти желаемыми должны быть несколько альтернатив. Значит, неопределенность должна состоять из желаемых реальных альтернатив (*plurality conditions*). И следовательно, реальные альтернативы должны существовать.

Шаги с первого по третий уже встречались выше. Новым является определение волеустанавливающего действия. Кейн объясняет: «Действия без достаточных причин будут действиями, в которых волевые установки агента не были еще однозначно определены до момента совершения действия. Именно в ходе совершения самих этих действий агент установит свою волю окончательно» [Kane 2007, 20]. Эти действия отличаются от действий с реальными возможными альтернативами в «примерах Остина». Действия в этих примерах не являются волеустанавливающими, так как воля агентов в них установлена заранее. Так, игрок в гольф до момента неопределенности уже определился в намерении попасть в лунку, убийца – убить премьер-министра, а Остин – выпить кофе без сливок. Поэтому их исход был просто слу-

чайностью. И примеры в стиле Остина не являются иллюстрациями волеустанавливающих действий.

Другое дело, если бы желания агентов в примерах не были однозначными, если бы у них были конфликтующие желания. Волеустанавливающие действия – это действия в ситуации выбора между двумя и более конкурирующими желаниями, когда выбор этот не ясен вплоть до момента осуществления действия. Тогда именно действие устанавливает волю. Так, если бы Остин хотел и кофе без сливок, и капучино, а убийца – и убить премьер-министра, и завязать со своей криминальной карьерой, не причиняя никому вреда, и эти желания конфликтовали бы вплоть до самого действия – вот тогда это были бы примеры устанавливающих волю действий.

Из этих же новых примеров следует и то, почему у волеустанавливающих действий должны быть желаемые альтернативы. Ведь только тогда любой из желаемых исходов не будет случайным. Для остановки бесконечного каузального регресса нужны действия с недостаточными причинами. У этих недетерминированных действий должны быть конфликтующие основания. Эти конфликтующие основания – желания альтернативных действий. А если альтернативы желаемы, они получаются и реальными, так как в недетерминированной ситуации любой из сценариев может реализоваться. Таким образом, Кейн из Принципа ответственности в конечном счете выводит сначала индетерминизм, а затем

– Принцип альтернативных возможностей. Эти два вывода он называет *двойным регрессом* свободы воли. Но допустим, что ему удастся провести этот двойной регресс. Допустим, его аргументы справедливы, и Кейн поднялся на вершину инкомпатибилизма. Что дальше, ведь, как он сам утверждал, это только половина пути?

3. Свобода воли и индетерминизм совместимы

А. Контрпримеры к классическому аргументу случайности

Индетерминистические миры могут быть устроены различным образом. Можно представить индетерминистический мир как череду случайных событий. В таком мире объекты будут возникать, пропадать и вновь появляться непредсказуемо в совершенно разных местах и конфигурациях. История такого мира будет представлять хаос, ленту из случайно подобранных кадров, персонажи которой перемешиваются, исчезают, преобразуются и превращаются один в другого. В подобной ситуации нельзя будет уловить сюжетную линию или предвидеть события. Нельзя будет предвидеть будущее даже на мгновение вперед. И объяснить наличное состояние также будет невозможно. Ведь в таком мире не будет

регулярностей, рациональности и законосообразности. Конечно, в таком мире не будет места ни контролю, ни ответственности. Но это не значит, что у сторонника либертарианского подхода нет выхода. Описанная ситуация, в которой все события случайны, – лишь карикатура на либертарианскую альтернативу мира с индетерминизмом. Существуют и другие, более осмысленные варианты.

Кейн конструирует вариант мира, где индетерминизм существует в определенных рамках, где неопределенность возникает только эпизодически. При этом эти эпизодические индетерминированные события не являются беспричинными. Просто причины не с необходимостью определяют действия, а с некоторой вероятностью. Это – вероятностная каузальность. При этом в случае, если действие не случается, случается другое, но совсем не любое. Возможными событиями являются не любые случайные события, а ограниченное число альтернатив. В таких мирах нет хаоса или просто произвольного стечения событий. Они вполне регулярны и предсказуемы. Именно такое устройство мира представляется необходимым Кейну для существования свободы воли и ответственности. За иллюстрациями далеко ходить не придется. Подходящие случаи индетерминированных, но контролируемых, предсказуемых и объясняемых событий уже встречались в нашей работе.

Речь идет о примерах «в стиле Остина». Кейн в данном случае использует их в качестве контрпримеров к Аргументу

случайности. Напомню один из них. Наемный убийца совершает покушение на премьер-министра. Он направляет снайперскую винтовку на жертву. Но в его движениях остается место неопределенности, связанной с квантовыми эффектами в его мозге. И от этой неопределенности зависит, будет выстрел точным или нет. Вот происходит выстрел, пуля попадает в государственного чиновника, и премьер-министр погибает, а убийцу хватают агенты спецслужб. В этой ситуации убийца совершает намеренное действие, даже несмотря на неопределенность в цепочке событий. И он не сможет объяснить свою непричастность тем, что между его решением выстрелить и выстрелом имели место индетерминированные нейронные процессы. Ведь его намерения были уже определены. Случайность могла ему только помешать. Для того чтобы вменить вину, есть достаточные основания – *mens rea* и *actus reus*. Убийца действовал преднамеренно. И он виноват, он несет ответственность. Индетерминизм в данном случае, очевидно, сочетается с ответственностью.

Ну а вот другой пример. Математик пытается решить сложное уравнение. Его намерение определено, но результат вычислений зависит от квантовой неопределенности в активации нейронных сетей его мозга. Допустим, он решает уравнение. Можно ли считать это действие намеренным и объяснимым? По всей видимости, да. Ведь он старался это сделать, и сделал это. Неопределенность могла бы быть лишь помехой в работе его мозга, как шум, помехи, меша-

ющие слушать радиопередачу. Результат его действия объясним, рационален и отчасти предсказуем. Это – также случай совмещения индетерминизма с контролем. Эти иллюстрации позволяют Кейну опровергнуть классический аргумент Юма и Принцип случайности. Но они не достаточны, чтобы на них построить позитивную либертарианскую теорию. Примеры Остина не дают возможность личности реализовать свободу выбора, свободу воли.

Математик и убийца в приведенных выше примерах отчасти контролируют с помощью своих желаний свои действия. Но, во-первых, степень этого контроля даже ниже, чем в случае полной детерминации, ведь на пути реализации их действий лежит неопределенность. И во-вторых, что еще более важно, математик и убийца не контролируют свой выбор. Для того чтобы построить успешную либертарианскую теорию, необходимо показать, что существует возможность не только сохранить, но и расширить контроль личности над своими состояниями. И этот контроль должен распространяться на сам выбор. С этой целью Кейн изобретает собственные примеры. Они должны удовлетворить всем описанным выше требованиям.

В. Индетерминизм и расширенный контроль

Новые гипотетические ситуации сконструированы таким

образом, что неопределенным оказывается не результат действия, а решение личности. То есть зона неопределенности перенесена с периферии действия ближе к центру; из области, где конкретное решение уже выполняется, в область, где только принимается решение. Из этих гипотетических ситуаций – назовем их «примерами в стиле Кейна» – я приведу две: «Ограбление» и «Выбор Ральфа».

*Ограбление*³³. Молодой человек во время прогулки в парке замечает, что происходит ограбление: преступники отбирают сумку у прохожей. Перед молодым человеком встает дилемма: с одной стороны, моральный долг призывает его вмешаться в конфликтную ситуацию, заступиться за женщину. С другой стороны, его пугает, что попытка защитить женщину может закончиться неприятностями для него самого. Молодому человеку нужно решиться: пройти мимо или попытаться помочь. Альтернативы действий продиктованы конфликтующими мотивами. Молодому человеку приходится выбирать между эгоцентрическими интересами и моральными альтруистическими установками. И этот выбор сопровождается внутренней борьбой. Исход борьбы мотивов и сопряженных с ними процессов в мозге, контролирующих поведение молодого человека, индетерминирован. В итоге молодой человек делает выбор – вступить за прохожую, и совершает поступок в соответствии со своими моральными установками.

³³ Аналогичный пример приводит Р. Кейн в книге [Kane 1996, 126].

*Выбор Ральфа*³⁴. Ральф совершает выбор жизненного пути: он раздумывает, оставаться ли ему в Мейберри или переезжать в Нью-Йорк. Мотивами переехать в Нью-Йорк служат его желания играть за футбольную команду «Гиганты» и быть звездой в бродвейском шоу. Мотивами остаться служат его желания взять замуж Робби Анну и управлять местным рестораном «Дер Виннершницель». Альтернативы несовместимы между собой, но в одинаковой степени желаемы для Ральфа. Выбор альтернативы сопровождается внутренней борьбой, исход которой индетерминирован. В результате индетерминистического процесса Ральф принимает спорное решение ехать в Нью-Йорк. Он *просто решает* ехать в Нью-Йорк.

По мнению Кейна, в этих двух примерах реализуется свободный выбор личности и совершаются волеустанавливающие действия. Это становится возможным из-за конфликта между двумя или более мотивами равной силы. Разрешение конфликта между ними индетерминировано. То есть перед агентом возникает возможность альтернативного поведения, реальная возможность. В ней реализуется расширенный контроль. Философ считает, что именно в этот момент обрывается каузальная цепочка, разрушается бесконечный регресс причин и действий. Агент становится отчасти независим от своей каузальной истории, как бы воспаряет над

³⁴ Подобный пример приводит М. Балагер, автор схожей с кейновской либертарианской позиции [Balaguer 2010, 72].

ней. В подобных ситуациях реализуется свобода воли и возникает фундамент для ответственности. Такие ситуации являются ключевыми в биографии агента. Благодаря им формируется характер личности, ее «добрая» или «злая» воля.

Волеустанавливающие действия определяют характер агента, который далее будет детерминирующим для многих последующих действий. И именно они оправдывают ответственность и за все последующие детерминированные действия, за действия, у которых уже не будет реальных альтернатив. Эти последующие действия будут результатом установок воли, определенных волеустанавливающими действиями. Накопление ВУД приводит к накоплению оснований для ответственности. Чем больше таких действий произведено, тем больше оснований считать человека ответственным за свое текущее состояние, тем больше оснований считать его последующие действия зависимыми именно от него, а его за них – *ответственным «в конечном счете»*. Так обстоит дело и с примером о Лютере. Кейн полагает, что Лютер действительно мог не иметь альтернатив, но он отвечал за свою позицию в силу сформировавших ее волеустанавливающих действий – тех самых мучительных размышлений и индетерминированного выбора в ситуациях неопределенности, о которых упоминалось выше.

Возможность осуществления волеустанавливающих действий в мозговых процессах, по мнению Кейна, не противоречит физике и нейрофизиологии. Об индетерминистиче-

ских процессах на квантовом уровне было уже сказано. Как я говорил выше, квантовая неопределенность сглаживается на макроуровне. Тем не менее Кейн считает возможным периодическое распространение этой неопределенности и на макроуровень. Эта возможность обосновывается им указанием на способность мозга к параллельным вычислениям. Широко известно, что мозг, в отличие от однопроцессорного компьютера, способен выполнять несколько когнитивных задач одновременно. Эта особенность мозга и создает механизм усиления неопределенности. Это происходит, когда есть внутренний конфликт, неопределенность в уме, пишет Кейн, когда напряженная работа мозга приводит к нарушению термодинамического равновесия – возникновению «хаотических процессов в мозге, которые делают его чувствительным к индетерминированным микропроцессам на нейронном уровне» [Kane 2013, 199]. Так Кейн укрепляет свою теорию с помощью современных нейрофизиологических представлений.

Но, допустим, нейрофизиологическая возможность ВУД доказана. Почему, тем не менее, мы должны все-таки верить интуициям Кейна в отношении этих случаев? Почему должны считать молодого человека в ситуации индетерминированного выбора ответственным за принимаемое решение? Почему эти ситуации демонстрируют расширенный контроль личности над своим поведением по отношению к случаям с полной детерминацией? У либертарианца есть два

основания считать так.

Во-первых, выбор, сделанный агентом, оценивается им самим как его собственный. Это – феноменологическое основание либертарианства Кейна. Хотя в выборе присутствует индетерминизм, агент признаёт, что выбор сделал он сам. Ведь он сам напряженно размышлял над альтернативами, и эти альтернативы были продиктованы его собственными ценностями, желаниями и убеждениями. Он не может пожаловаться, что выбор случился сам по себе и что это неожиданное для него решение. Дело было бы совсем иначе, если бы выбор пал на альтернативу, которую агент даже не обдумывал. Тогда, возможно, он бы не принял этот выбор, а удивился ему. Но в описанных ситуациях это был выбор самих агентов, по крайней мере из феноменологической перспективы.

Во-вторых, о выборе, описанном выше, нельзя сказать, что это выбор случайный. Когда говорят о случайном действии, говорят обычно об ошибке, недоразумении, произвольном, непреднамеренном событии. Но разве выбор вступить за прохожую был случайным, ошибочным недоразумением? Он был обоснованным, его можно полностью объяснить мотивами агента. Ситуация в целом похожа на ту, когда убийца, намеревающийся убить премьер-министра, попадает в цель. И не похожа на ту ситуацию, когда убийца ненароком промахивается. Ведь действия агента соответствуют его намерениям. Специфика ситуации с ВУД, однако,

в том, что намерениям соответствуют оба исхода, обе альтернативы. В обоих случаях агент успешно совершает намеренные действия, в обоих случаях ему удастся сделать то, что он хочет. Даже если бы молодой человек не бросился на помощь прохожей, а скрылся бы, – это тоже было бы успешным выполнением его замысла. В этом смысле он имеет контроль над обеими перспективами.

Кейн называет этот «расширенный контроль» *множественным волевым контролем*. Обладать «множественным волевым контролем» над набором альтернатив – значит иметь возможность реализовать любую из альтернатив, которую желаешь, в соответствии с собственными предпочтениями и мотивами, без внешнего принуждения [Kane 2013, 209–210]. Множественный волевой контроль над действиями является основанием для ответственности.

4. Возражения к теории Кейна

Теория Кейна продолжает традиции классического либертарианства. Ключевой идеей этой теории, как и теорий предшественников Кейна, является утверждение о том, что индетерминизм открывает перед агентом альтернативные возможности действий и прерывает бесконечный регресс каузальной ответственности. При этом американский философ значительно укрепляет либертарианские позиции, развивая аргументы этой теории и детализируя описания сво-

бодных действий. С помощью «случаев в стиле Остина» философу удастся показать, что индетерминизм не всегда обращивается случайностью. Основанием возможности агента контролировать свое поведение в индетерминистической ситуации, по Кейну, служит вероятностная каузальность. Согласно Кейну, только она оставляет агенту возможность множественного волевого контроля, необходимого для ответственности. Философ согласовывает свою теорию с современными представлениями в области квантовой физики и знаниями о параллельной архитектуре мозга, полученными современными нейрофизиологами. Все это, безусловно, является преимуществами его теории. Но у нее есть и существенные недостатки.

На мой взгляд, (1) Кейну не удастся показать практическую значимость индетерминистической свободы. И, что еще важнее, (2) ему не удастся продемонстрировать, как индетерминизм создает принципиальную возможность *ответственности в конечном счете*. Таким образом, с моей точки зрения, его теория не решает основную поставленную перед ней задачу, не дотягивает до стандартов, которые сама же задает. Это не говорит о несовместимости индетерминизма и свободы. Скорее это демонстрирует несостоятельность концепции *ответственности в конечном счете*, на которую опирается Кейн. «Ответственность в конечном счете» является недостижимой целью как в индетерминистическом, так и в детерминистическом мире. При этом, с моей точки зре-

ния, она избыточна в практическом отношении, для объяснения взаимоотношений между людьми. Сейчас я объясню, в чем, на мой взгляд, заключаются основные недостатки аргументации Кейна.

А. Практическая значимость и возможность волеустанавливающих действий

Кейн утверждает, что либертарианская концепция во многом согласуется с бытовыми интуициями: их он считает одной из опор своей теории. Философ также настаивает на том, что свобода проявляется по мере накопления волеустанавливающих действий, так как эти действия способствуют становлению агента как автономного источника собственного поведения. Из теории Кейна следует, что сомнения перед действием увеличивают степень вины. И что мы никого не можем признать свободным и ответственным до тех пор, пока не проведем детального исследования его или ее жизни и не обнаружим этих сомнений. Но, как мне кажется, эти соображения плохо согласуются между собой и, более того, ведут к практическим недостаткам его либертарианской теории.

Обыденная практика определения степени ответственности и свободы предполагает выявление замысла поступка. Поступок вызывает особенное осуждение, если он совершен осознанно и намеренно. Но практика приписывания ответ-

ственности не предполагает учета сходных с волеустанавливающими действиями ситуаций в биографии личности. Степень ответственности не кажется большей в связи с тем, что человек сомневался, как поступить. Наоборот, убежденный эгоист, жесткосердный преступник, никогда не задумывавшийся об интересах других людей, уверенно действующий вопреки их интересам, представляется достойным большего осуждения. Конечно, иногда нас все-таки может интересовать, сомневался ли преступник перед совершением преступления. Но этот интерес связан лишь с тем, что сомнения являются индикатором осознанности действий.

Сомнения принимаются в расчет при оценке действий и меры ответственности личности главным образом потому, что они выступают доказательствами вменяемости преступника, наличия у него злых намерений, умышленности действий. Если человек понимал, что делает, и не сомневался в своих действиях, то он несет еще большую ответственность. А из теории Кейна следует обратное. Чем сильнее были неосуществившиеся желания, чем дольше и больше раз имела место внутренняя борьба, тем, по его мнению, больше ответственности несет личность за выбранную альтернативу. Мне представляется это контринтуитивным. Более того, это ведет к значительным практическим трудностям.

Что, если у личности в автобиографии совсем не было волеустанавливающих действий? Что, если среди конкурирующих мотивов всегда был явный фаворит, и они не стал-

квивались в непримиримой и столь мучительной для агента борьбе? И какова вероятность, что волеустанавливающие действия существовали в истории каждого осуждаемого агента? Согласно теории Кейна, их существование абсолютно необходимо для ответственности. Но, если принять эту теорию, необходимо провести ревизию всех случаев осуждения и остановить практику. До тех пор пока наличие достаточного количества волеустанавливающих действий не доказано, не рационально считать агентов ответственными. Из идей Кейна следует необходимость довольно радикальных изменений в общении между людьми. Необходимость ревизии обыденных практик я считаю признаком практической неприменимости либертарианской теории Кейна. Впрочем, в ней есть, как мне кажется, и чисто теоретические проблемы.

В. Отсутствие ответственности в конечном счете в случае ВУД

Главной целью Кейна является демонстрация возможности *ответственности в конечном счете*. Этот тип ответственности предполагает ответственность за все причины события. В качестве парадигмы таких случаев Кейн приводит волеустанавливающие действия. По его мнению, такие действия реализуются в «случаях в стиле Кейна». Но, как мне кажется, эти случаи совсем не так однозначны, как хо-

тел бы представить их автор. В контексте либертарианской теории, после рассуждений Кейна о несовместимости ответственности и детерминизма, читатель может склониться к выводам, которые предпочитает философ. Можно допустить наличие ответственности за выбор в случаях волеустанавливающих действий. Но если абстрагироваться от контекста, «случаи в стиле Кейна» можно интерпретировать иначе. Можно вписать их в другой контекст и обратить внимание на другие детали. Тогда интуиции склонят нас в пользу отсутствия ответственности за принятое решение.

Так, в случае с «Ограблением» можно обратить внимание на то, что молодой человек в двух соседних возможных мирах – там, где он в момент t помогает прохожей, и там, где решает пройти мимо, – имеет совершенно одинаковые предшествующие состояния. Ничто в его убеждениях и желаниях, размышлениях и усилиях, характере и достоинствах не отличается в этих двух мирах. И тем не менее в одном случае он выходит из ситуации трусом, а в другой – героем. Как объяснить разницу в его поступках? А если объяснения Кейн предоставить не может, разве тогда исход – не случайность? Разве нельзя утверждать, что молодому человеку просто повезло в одном случае стать героем и не повезло в другом случае, когда он повел себя как трус? И следствием случайности станут и другие последующие его поступки, ведь именно случайность установит его характер.

Интерпретация «случаев в стиле Кейна» как ответствен-

ных действий зависит от их связи с «примерами в стиле Остина». Именно на их экстраполяции строит заключения Кейн. Но связь между этими типами случаев в отношении ответственности проблематична. Ведь зона индетерминизма в них расположена в разных местах, а это принципиально важно. В самом деле: существует несколько вариантов расположения зоны индетерминизма в функциональном устройстве агента. Во-первых, действие может начинаться индетерминированным образом. То есть альтернативы действий могут формироваться в ходе индетерминистического процесса. Во-вторых, выбор из набора альтернатив может осуществляться индетерминистически. И наконец, в-третьих, сама реализация действия может проходить с элементами индетерминизма. Кейн сначала приводит случаи «в стиле Остина», относящиеся к третьей категории, для доказательства совместимости индетерминизма с ответственностью. А потом делает экстраполяцию на вторую категорию, на «случаи в стиле Кейна». Но это, мне кажется, – неоправданный шаг. Ответственность, очевидно присутствующая в случаях первого типа, убывает в случаях, относящихся ко второму типу. Перенесение индетерминизма с этапа реализации действия к моменту принятия решений значительным образом снижает ответственность. Продемонстрировать это наглядно позволит следующая иллюстрация.

Предположим, студент решает математическую задачу, в которой есть один правильный ответ. В случае А он дает

один собственный ответ. В случае *B* он дает несколько ответов. При этом он не может сделать выбор между ними. В третьем случае, *C*, студент использует уже готовые варианты ответов и выбирает один из них. В каком случае он несет большую ответственность за решение? В случае *A*, как мне кажется, степень ответственности самая высокая. Либо решение правильное, и студент должен получить высокую оценку, либо решение неправильное, и оценка должна быть низкая. В случае *B* ответственность, кажется, меньше: позиция студента неопределенна. Оценить его работу сложно. Так как даже если среди его решений есть правильное, он не определился до конца с выбором. Тем не менее, если среди его ответов нет правильного, можно сказать, что его решение точно неверное, и он заслуживает отрицательной оценки. А как быть в третьем случае?

Мне кажется, что в случае *C* ответственность студента за решение еще меньше. Если ответ правильный, нельзя сказать, насколько студент на это повлиял – он мог выбрать его наугад. Аналогично, если он дал неправильный ответ, он может быть к этому непричастным, если среди вариантов ответа не было верного. Думаю, иллюстрация показывает, что передвижение зоны индетерминизма от реализации к принятию решения влечет убывание ответственности. Так, в первом случае выбор и формулировка решения всецело принадлежат самому студенту, и потому он несет полную ответственность. Во втором случае, сформулировав решения, но

не осуществив выбор одного из них, студент несет частичную ответственность. В третьем случае, когда сами варианты решений представлены студенту уже готовыми, ответственности еще меньше. Из этого следует, что оснований распространять выводы, относящиеся к случаям «в стиле Остина», на случаи «в стиле Кейна» у философа нет.

Важно отметить, что я здесь не утверждаю, что в случаях «в стиле Кейна» ответственность совсем отсутствует. Некоторая степень ответственности, по моему мнению, все же есть. Но она меньше, чем в случаях «в стиле Остина», и эта степень уж точно не соответствует требованиям философа к ответственности в конечном счете. А из этого следует, что накопление ВУД в личной истории агентов ведет не к повышению их ответственности за действия, а к снижению. Это прямо противоречит утверждениям Кейна, ведь, по его мнению, чем больше ВУД в истории человека, тем выше степень его ответственности *в конечном счете*. Свою интуицию я попытаюсь обосновать с помощью собственного примера. Мне кажется, этот тип примеров можно использовать для критики теории Кейна.

Предположим, охраннику предложили за вознаграждение убить премьер-министра. Размышляя об этом предложении, охранник испытывает сомнения. С одной стороны, к убийству его подталкивают корыстные мотивы. Он представляет, как заказчики выплатят ему огромную премию, благодаря которой он сможет обеспечить безбедное существование

своим детям. С другой стороны, его мучает совесть. Он понимает, что в результате его действий погибнет невинный человек. Это – случай ВУД, поскольку выбор не детерминирован. Но допустим, что, неведомо для него, за состоянием его мозга следит заказчик преступления. Некоторое время назад он установил в мозг охранника радиопередатчики, которые позволяют наблюдать за его состояниями и влиять на них. Передатчики этого устройства не очень надежны: есть лишь некоторая вероятность, что с их помощью удастся определить выбор охранника, то есть в их действии имеется индетерминизм. Обнаружив, что охранник еще сомневается и покушение может не состояться, заказчик пытается вмешаться в работу его мозга. Устройство срабатывает, и выбор охранника определяется в пользу покушения. Несет ли охранник ответственность в конечном счете за убийство? Является ли это решение контролируемым самим агентом, или оно продиктовано чем-то внешним?

Думаю, здесь ответственность охранника можно считать ограниченной. В каком-то смысле он виноват, что допустил возможность выбора варианта с преступлением. При этом он реально мог поступить иначе. Однако исход в итоге был во многом определен вмешательством заказчика покушения. Значит, контроль и ответственность самого охранника редуцированы, и он не обладает ответственностью в конечном счете. Хотя при этом выполнены условия множественного волевого контроля. С феноменологической точки зре-

ния охранник будет считать это своим выбором, и выбор нельзя назвать случайным.

Сторонник Кейна мог бы возразить: но в описанном выше случае очевидно вмешательство стороннего агента! Это случай манипуляции. Естественно, сложится впечатление, что контроль редуцирован. Но на это я отвечу: да, конечно, в этом конкретном случае есть вмешательство стороннего агента. И это – действительно манипуляция. Но заказчика с его радиопередатчиком могло бы не быть, а результат остался бы тем же. Просто роль заказчика исполнил бы случай. Случай тоже может быть манипулятором. Стороннику Кейна нужно показать, чем отличается ситуация, когда заказчик с помощью индетерминированного процесса управляет решением охранника, от ситуации, когда решением охранника управляет случай. А до тех пор, пока он этого не сделал, можно утверждать, что, во-первых, индетерминированность событий в случае ВУД уменьшает контроль личности над действием и, соответственно, уменьшает ответственность. И во-вторых, наличие ВУД в биографии личности не обеспечивает возможность ответственности в конечном счете. С учетом этих обстоятельств стоит рассмотреть другую альтернативу решения проблемы свободы воли и моральной ответственности – концепцию младшего современника Р. Кейна, Д. Перебума.

III. Свобода воли и моральная ответственность в жестком инкомпатибилизме Д. Перебума

По мнению Р. Кейна, моральная ответственность несовместима с детерминизмом, и Д. Перебум с этим согласен. Поэтому позиции обоих философов являются вариантами инкомпатибилизма. При этом Кейн считает, что свобода воли и моральная ответственность совместимы с индетерминизмом, а по мнению Перебума – нет. Моральная ответственность и свобода воли несовместимы как с детерминизмом, так и с индетерминизмом, считает Перебум. Поэтому его позиция называется жестким инкомпатибилизмом. Согласно этой теории, в любом из миров, где каузальность связывает события необходимой или вероятностной связью, ответственность исключена. Ответственность, по его мнению, существовала бы, только если бы каузальность связывала субстанции, не детерминированные другими причинами, и события. Но такое положение дел, вероятно, противоречит науке. Следовательно, в нашем мире, скорее всего, свободы и моральной ответственности нет.

Скептический вывод Перебума относительно свободы, однако, не подрывает основания морали и смысла жизни. По мнению философа, он только способствует эволюции

этих представлений. Классическая концепция ответственности во многом опирается на идею базовой заслуги [Pereboom 2014b, 2], но это не единственно возможная опора. Можно отказаться от идеи заслуги и сохранить некоторый вариант ответственности. В соответствии с этой обновленной идеей ответственности, поощрять людей будет все равно оправданно, так как поощрение способствует распространению положительных примеров в обществе. Изменениям подвергнется практика наказаний – она приблизится к практике «лечения» и «карантина». Такого рода наказания сохранят свою основную функцию: они будут оберегать общество от криминального поведения и препятствовать его распространению. При этом они будут более гуманны в отношении преступников. «Эти предложения не исключают из нашей практики применения моральной ответственности даже в криминологическом аспекте. Скорее они приведут к пересмотру практики с учетом уже существующих правил» [Pereboom 2014b, 174].

В этих утверждениях – суть позиции жесткого инкомпатибилизма Перебума. Теперь я рассмотрю основные аргументы в их пользу. Сначала стоит обратиться к аргументам Перебума против совместимости свободы воли и индетерминизма, то есть против либертарианства, в частности против позиции, которую отстаивает Кейн. И далее я обращусь к аргументам против совместимости свободы воли и ответственности с детерминизмом, то есть к доводам Перебума против

компатибилизма. Моя собственная позиция и критические замечания в отношении позиции Перебума и инкомпатибилизма в целом будут представлены в последней, четвертой части этой главы.

1. Свобода воли и индетерминизм несовместимы

А. Индетерминизм не расширяет контроль

Кейн считает, что индетерминизм становится основанием моральной ответственности, когда конфликтующие мотивы достигают равной силы и агент вынужден принимать одно из решений («случай в стиле Кейна»). В такой ситуации перед агентом раскрываются альтернативы. В силу индетерминизма агент может поступить так или иначе, но выбор при этом не перестает быть выбором самого агента. Контроль над выбором в таких ситуациях не теряется. Напомню, Кейн представляет два основания считать так. Во-первых, решение объективно не перестает быть желаемым, намеренным, понятным, целенаправленным. Оно обуславливается мотивами агента и является результатом его усилий. Нельзя назвать это решение случайным или ошибочным. Во-вторых, агент субъективно считает эти решения собственными. То есть с феноменологической точки зрения они представля-

ются ему контролируруемыми решениями. Но эти основания присутствуют и в арсенале компатибилиста, – считает Перебум. Они существовали бы даже в том случае, если бы «трудный выбор» агента происходил в детерминистическом мире: из феноменологической перспективы детерминированный выбор также представлялся бы агенту не случайным, а намеренным, желанным, ожидаемым. И объективно этот выбор соответствовал бы мотивам агента. Тогда в чем тут преимущество индетерминированного выбора?

Кейн считает, что преимущество заключается в доступности альтернативных возможностей. По его мнению, альтернативные возможности создают условия для самоопределения агента. Они позволяют осуществить множественный волевой контроль в случае волеустанавливающих действий. А такой вид контроля не может присутствовать в детерминистическом мире. Перебум, однако, не видит в индетерминизме и альтернативных возможностях особой ценности. С его точки зрения, они не дают агенту никакой дополнительной свободы. По его мнению, агент одинаково ответствен в случаях как с, так и без альтернативных возможностей. Позиция Перебума здесь основывается на уже упомянутых выше примерах Франкфурта. Философ выступает защитником этих примеров, точнее, их модифицированных, усовершенствованных версий, поскольку исходные примеры Франкфурта не лишены, по мнению Перебума, существенных недостатков.

а) Примеры «в стиле Франкфурта» и проблема контролера

Примеры Франкфурта, как я уже писал, для современных аналитиков являются традиционным способом критики Принципа альтернативных возможностей. Они должны показать, что отсутствие альтернатив может сочетаться с наличием ответственности. Но у них есть недостаток – так называемая проблема контролера. Она заставляет усомниться в чистоте предлагаемого мысленного эксперимента. Продемонстрировать ее проще всего на классическом примере Франкфурта, в котором Блэк (контролер) вживляет Джонсу в мозг электроды, чтобы при необходимости заставить его сделать А. Но в этом примере Джонс совершает А по собственному желанию, без вмешательства Блэка, и, по всей видимости, сам несет ответственность за А. Ведь Блэк так ничего и не предпринял.

На первый взгляд, этот классический пример Франкфурта вполне убедителен. Может показаться, что он достигает поставленных автором целей: демонстрирует ситуацию, в которой у агента отсутствуют альтернативы, но имеется ответственность. Однако, как и в случае с большинством других мысленных экспериментов, здесь многое зависит от деталей, прояснение которых может оказаться принципиальным. Необходимо определить, как именно контролер Блэк узнает, что ему не требуется вмешиваться. Что должно произойти, чтобы он точно понял, что не нужно действовать?

Ведь если Блэк не будет иметь ориентиров, воображаемая ситуация невозможна.

В исходной статье «Альтернативные возможности и моральная ответственность» Франкфурт предлагает несколько вариантов того, как Блэк мог бы влиять на выбор Джонса: это и угрозы, и специальное снадобье, и гипноз, и радиоуправляемые электроды, воздействующие на нейронные сети Джонса. Но он почти ничего не говорит о том, на какой сигнал Блэк должен ориентироваться. Франкфурт ограничивается только уверением в том, что Блэк – «превосходный эксперт в этих вопросах» [Frankfurt 2013, 144]. Остается додумывать самим. Можно предположить, что ориентиром для действий Блэка могло бы служить состояние мозга Джонса, предшествующее решению. Так, предположим, что выбору в пользу действия A в момент t_2 предшествует состояние NA в момент t_1 . Если Блэк видит, что в момент t_1 мозг Джонса находится в этом состоянии, он заключает, что Джонс сам сделает A , и не вмешивается. Если же в t_1 у Джонса – какое-то другое состояние, Блэк будет действовать. Но теперь неясно, действительно ли у Джонса нет альтернатив. Удастся ли Франкфурту действительно совместить ответственность и отсутствие альтернатив у агента?

Если принять, что в мысленном эксперименте Франкфурта Блэку требуется некоторый сигнал, чтобы понять, когда начать действовать, необходимо принять и возможность различных сигналов. Так, Джонс, очевидно, в момент t_1 может

иметь состояние NA , а может и не иметь его. Несмотря на то что Джонс не может поступить иначе, он может оказаться в том или другом предшествующем состоянии. Назовем это *проблесками свободы*. Сторонник альтернативных возможностей как условия моральной ответственности может теперь зацепиться за это. И настаивать на том, что именно «проблески свободы» в примерах Франкфурта оправдывают интуиции о моральной ответственности агента.

Но, думаю, это слишком ненадежная опора. Ведь речь идет только о непреднамеренных неосознаваемых промежуточных состояниях. Эти состояния равносильны рефлексам, случайным вздрагиваниям, оговоркам, тикам и т. п. Они не выражают осознанный выбор агента. Вряд ли возможность подобных произвольных событий может служить основанием ответственности. Абсурдно же утверждать, что основание ответственности за убийство заключается в возможности промахнуться. Впрочем, защита «проблесков свободы» как альтернативных возможностей, на которых основывается моральная ответственность, – не единственный путь опровергнуть примеры в «стиле Франкфурта». Философы Р. Кейн, Д. Вайдеркер и К. Жине предлагают иную линию критики. Они считают, что в описываемых Франкфуртом ситуациях можно отыскать не просто «проблески свободы», но реальные альтернативы.

Доказательство этого начинается с определения, в каком мире реализуются сценарии Франкфурта. Сначала предпо-

ложим, что это – детерминистический мир. В детерминистическом мире по предшествующему сигналу Блэк может достоверно определить, нужно ли ему действовать. Но проблема в том, что детерминизм, согласно представлениям инкомпатибилизма, исключает моральную ответственность. В таком мире, согласно аргументу, представленному выше, агент не является «ответственным в конечном счете». Настаивать на моральной ответственности агента в этом случае – значит просто бездоказательно отрицать позицию инкомпатибилизма.

Другое дело, если мир, в котором реализуются описываемые Франкфуртом ситуации, индетерминистический. В таком мире, по мнению инкомпатибилиста-либертарианца, ответственность может существовать. Но тогда и возникает сомнение в отсутствии у агента альтернативных возможностей. Представим, что имеет место воле-устанавливающее действие. Между предварительным сигналом о состояниях агента в момент t_1 и его выбором в момент t_2 отсутствует детерминистическая связь. Но тогда у контролера-манипулятора не будет достаточного основания вмешиваться. Если он вмешается, то агент ответственности, очевидно, нести не будет. Выбор за него осуществит контролер. А если контролер не станет вмешиваться, у агента в момент t_2 останется возможность поступить иначе. Ведь индетерминизм предполагает, что при заданных начальных условиях с определенной вероятностью могут произойти различные собы-

тия, а значит, есть и вероятность, что агент осуществит другое действие. Таким образом, без вмешательства контролера агент имеет альтернативные возможности и несет ответственность.

б) Пример Перебума и решение проблемы контролера

Перед защитником примеров «в стиле Франкфурта» встанет дилемма: либо в примере неочевидно наличие моральной ответственности агента, либо неочевидно отсутствие альтернативных возможностей. Назовем это *дилеммой контролера* в примерах в стиле Франкфурта. «Если агент ответственен в конечном счете, так как контролер не вмешался, агент мог поступить иначе; но если агент *не мог* поступить иначе, потому что контролер вмешался, тогда контролер ответственен в конечном счете, а не агент» [Kane 1996, 142–143]. Без решения этой дилеммы опровергнуть Принцип альтернативных возможностей не удастся. Чтобы ее решить, требуется модификация исходного примера Франкфурта. И такую модификацию предлагает Д. Перебум. Его пример «в стиле Франкфурта» очень изобретателен. Но автор убеждает, что взят этот пример частично из собственного опыта.

Пример 3. Уклонение от налога [Pereboom 2014b, 15]. Джо размышляет о том, воспользоваться ли ему налоговой льготой в связи с дорогостоящей покупкой. Он понимает, что в данном случае льгота не предусмотрена законом. Но, скорее всего, его ходатайство никто не оспорит. В любом случае он сможет оправдаться тем, что плохо разбирается

в налоговом законодательстве. У Джо есть сильное желание воспользоваться льготой и не платить налог. Принятие эгоистических решений, даже если они ведут к нарушению некоторых законов, для Джо приемлемо. Единственное, что могло бы побудить Джо принять решение не пользоваться льготой и заплатить все налоги, – определенный уровень внимания к моральным доводам.

Определенный уровень внимания к моральным доводам есть *необходимое, но не достаточное* условие выбора в пользу уплаты налога. Если Джо достигает определенного уровня внимания к моральным доводам, он, обладая либертарианской свободой воли, может сделать выбор как в пользу уплаты налога, так и в пользу уклонения от него. Блэк хочет, чтобы Джо уклонился от уплаты налогов. Поэтому втайне от Джо он установил в его мозг устройство, которое в случае достижения определенного уровня внимания к моральным доводам автоматически вынудит Джо выбрать уклонение от налогов. Джо размышляет над решением, он не достигает необходимого уровня внимания к моральным доводам и выбирает уклонение от налогов. Устройство остается бездейственным. Вероятно, в такой ситуации Джо несет полную ответственность за свой выбор.

Схема на *рис. 4* отражает обстановку мысленного эксперимента Перебума. Ромбиками обозначены состояния агента, стрелками с пунктиром – индетерминистические связи, а сплошными линиями – детерминистические. Толстой чер-

ной стрелкой обозначен случай интервенции контролера.

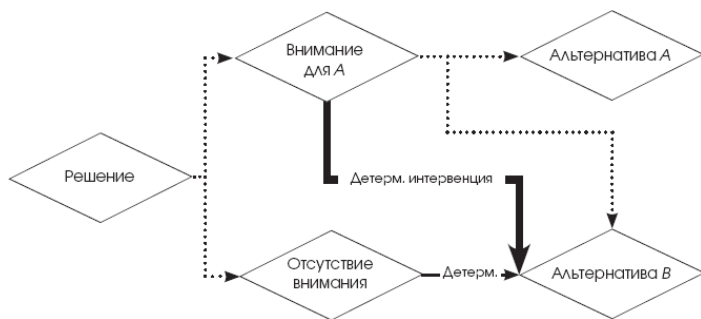


Рис. 4. Схема примера Д. Перебума «Уклонение от налога»

Согласно описанной ситуации, агент находится в мире, где есть индетерминированные события. В силу неопределенности во время принятия решения агенту доступны две альтернативы: сконцентрироваться на моральных доводах (Внимание для А) или нет (Отсутствие внимания). Концентрация внимания является необходимым условием принятия решения заплатить все налоги (Альтернатива А). Но это – не достаточное условие. Даже с учетом моральных доводов агент в силу неопределенности все еще может принять любое из двух решений (доступность альтернатив обозначена пунктиром). Однако для контролера это – уже достаточное основание вмешаться, ведь на этом этапе возникает риск, что Джо все-таки решит выплатить налог. Зато пренебрежение

моральными доводами со стороны Джо – верный признак того, что все само пойдет так, как надо Блэку, и ему не придется вмешиваться. В эксперименте есть надежный сигнал для вмешательства контролера и неопределенность, необходимая для либертарианства.

Теперь мы можем протестировать работу мысленного эксперимента. Проиграем воображаемую ситуацию. Джо задумывается о выборе. Через некоторое время он начинает концентрироваться на моральных доводах. Тут же вмешивается Блэк. Он заставляет Джо принять решение в пользу налогового вычета. Ответственности у Джо нет, и нет альтернатив. Воспроизведем эксперимент еще раз. Джо опять концентрируется на моральных доводах, и опять Блэк вмешивается, заставляя первого делать все тот же эгоистический выбор. У Джо просто нет шансов принять иное решение. Сколько бы раз эксперимент ни разворачивался по этому сценарию, итог будет одинаковым. А что будет, если реализуется другой сценарий? Предположим, Джо опускает моральные соображения и сам выбирает уклонение от налогов. Тогда Блэк остается бездействовать. Решение Джо индетерминировано, и, следовательно, даже по мнению либертарианца это будет ответственное решение. При этом у Джо нет реальной альтернативы выплатить налог. Значит ли это, что *дилемма контролера* преодолена?

У противника франкфуртских примеров есть еще один доступный шаг – показать, что у Джо все-таки есть какие-то

альтернативы. Хотя Джо не может не уклониться от налога, ему все же доступны разные варианты поведения. Он может сконцентрироваться или не сконцентрироваться на моральных соображениях. Это и есть альтернативы Джо. Противник франкфуртских примеров будет доказывать, что именно наличие этих альтернатив составляет основание ответственности. Если это так, то обстановка мысленного эксперимента не соответствует задачам. Она не позволяет показать совместимость ответственности и отсутствия альтернатив.

Перебум с этим выводом не согласен. Он доказывает, что альтернативы, доступные Джо в его обновленном примере, не являются *релевантными и существенными (robust)*, чтобы служить основанием ответственности. Они – не более чем «проблески свободы». Они не удовлетворяют критериям существенных релевантных альтернатив. Чтобы альтернатива была существенной и релевантной, она, во-первых, должна быть осознанной. Ведь абсурдно утверждать, что основание ответственности за поступок лежит в возможности случайно его не совершить. Агент должен сам желать совершить или не совершить некоторый поступок. Во-вторых, альтернативный поступок должен вести к другим последствиям: в результате выбора этой второй альтернативы агент не должен нести того же морального бремени, что и в случае выбора первой. Иначе в чем заключается ее альтернативность, если она приводит к тем же последствиям? В-третьих, релевантная и существенная альтернатива должна удовлетворять

эпистемическим критериям. Агент должен хотя бы отчасти понимать, что выбор в пользу второй альтернативы избавит его от той моральной ответственности, которая наступит в первом случае.

Перебум приводит пример, который иллюстрирует работу этих трех критериев. Назовем его «Кофе налогоплательщика». Предположим, Джо может как заплатить, так и не заплатить налог. Единственным случаем, когда Джо заплатит налог, является случай, когда он перед принятием этого решения выпьет чашку кофе. В чашку без его ведома добавлено вещество, заставляющее его действовать в соответствии с налоговым кодексом. Во всех остальных случаях Джо налог не заплатит. Итак, Джо не пьет кофе и не платит налог. Является ли единственная имевшаяся альтернатива основанием считать, что Джо ответственен за свое фактическое действие? Альтернатива Джо представляла (1) сознательное действие, (2) действие, ведущее к соблюдению им закона. Джо мог выпить кофе осознанно. И результатом этого действия было бы выполнение налоговых требований. Но это была нерелевантная и несущественная альтернатива. Ведь Джо не мог знать, что это действие приведет к другому действию, к выплате налога. Эта альтернатива не удовлетворяет третьему критерию существенности и релевантности. Следовательно, она не может служить основанием для наличия ответственности.

Альтернативой такого же рода является и возможность

Джо в примере «Уклонение от налогов» сконцентрироваться на моральных доводах, считает Перебум. Джо не знал, что эти размышления могут привести его к уплате налогов и другим моральным последствиям. Поэтому концентрация внимания не является в данном случае релевантной и существенной альтернативой. И, соответственно, она не является основанием для возникновения моральной ответственности. Если эти рассуждения верны, дилемма контролера решена. И из примеров в стиле Франкфурта можно сделать вывод, что основанием ответственности служит не наличие альтернатив, а способность агента быть «причиной в конечном счете». «Может статься, что альтернативные возможности – не релевантные и не существенные – являются следствиями того, что агент выступает источником своих действий... Однако такие альтернативные возможности не будут играть первостепенной роли при объяснении моральной ответственности агента. Скорее они будут следствием первостепенного фактора: того, что агент является источником своих действий...» [Pereboom 2013a, 431]. Таким образом, Перебум, в отличие от Кейна, считает, что альтернативные возможности не являются существенными для ответственности и не добавляют контроля агенту. И в этом – одно из важных отличий теорий Кейна и Перебума. Но самое принципиальное их отличие – в том, что Перебум считает свободу воли и индетерминизм несовместимыми. В пользу этого тезиса он приводит Аргумент исчезающего агента.

В. Аргумент исчезающего агента

Перебум относит этот свой авторский аргумент, Аргумент исчезающего агента, к семейству Аргументов случайности, о которых речь уже шла выше. Однако, в отличие от классического аргумента Юма, это рассуждение нацелено против современных вариантов либертарианства. В нашем тексте это уже вторая модификация аргумента. Вывод этого рассуждения касается ситуации, парадигмальной для современных сторонников либертарианства, и в частности для Кейна. Он опровергает наличие достаточного контроля в случаях волеустанавливающих действий.

Представим, что агент делает выбор между двумя альтернативами поступков. Моральные соображения склоняют его выбрать первую альтернативу, а эгоистические соображения – вторую. Альтернативы действий исключают друг друга, а мотивы в пользу каждой из них обладают одинаковой силой для агента. Если считать, что только события имеют каузальную силу, и принять условие, что релевантные antecedentes события, в которых участвует агент, не устанавливают однозначно (*settle*), какое из решений будет принято, но лишь определяют, что каждое из решений может быть принято с вероятностью, равной 0,5, то ничто не устанавливает, какое решение будет принято. Соответственно, ни сам агент, ни что-либо, относящееся к агенту, не устанавливает, какое

решение будет принято. Поэтому агент не имеет контроля, необходимого для моральной ответственности в смысле базовой заслуги (*basic desert*). Таково рассуждение

Перебума. К нему я хочу сделать три важных пояснения.

а) Мишенью Аргумента является не существование агента, а существование моральной ответственности

Название, которое автор дал аргументу, может сбить с толку: может показаться, что, согласно рассуждению Перебума, в индетерминированной предшествующими событиями ситуации неизбежно *исчезает агент* и что именно это исчезновение является выводом аргумента. Тогда для того, чтобы доказать, что индетерминизм ведет к отсутствию моральной ответственности, нужно было бы указать, что исчезновение агента, или исчезновение *того самого* агента, что совершил выбор, есть основание для исчезновения ответственности. Но мишенью для критики в данном аргументе служит все-таки не существование агента, а наличие у него расширенного контроля. Перебум считает, что в условиях, когда мотивация агента совершить альтернативные поступки равносильна, *теряется именно контроль*. Поэтому правильней было бы назвать его Аргументом исчезающего контроля. «Выявляемая проблема состоит в том, что, поскольку агенты, согласно событийно-каузальному либертарианству, не способны устанавливать, какое из решений будет принято, они не могут исполнять такую роль по отношению к действию, которая обеспечила бы контроль, требуемый этим ти-

пом моральной ответственности» [Pereboom 2014b, 32]. Индетерминизм в ситуации спорного выбора подрывает контроль агента, но без контроля выбор становится случайным, а за случайный выбор ответственность не наступает, по крайней мере в смысле *базовой заслуги*.

б) Аргумент критикует возможность не всякой моральной ответственности в условиях индетерминизма, а только возможность ответственности в смысле базовой заслуги

Перебум считает, что только ответственность в смысле базовой заслуги проблематична в данном случае и в дебатах о свободе воли в целом. Другие типы ответственности могут быть совмещены даже с вынужденными действиями или действиями случайными. Базовая заслуга предполагает приписывание ответственности агенту исходя лишь из того, что он совершил действие и понимал его моральное значение, без учета соображений о будущих событиях. «Чтобы агент был морально ответственным за действие в этом смысле, оно должно принадлежать ему таким образом, чтобы он заслуживал порицания, если понимал, что оно было неправильным с моральной точки зрения, и заслуживал похвалу, если понимал, что оно было морально безупречным. Заслуга в данном случае является базовой в том смысле, что агент заслуживает порицание или похвалу только потому, что совершил этот поступок, понимая его моральный статус, а, например, не в силу одних только соображений консеквенци-

алистского или конвенционального характера. Такая характеристика оставляет возможность считать агента морально ответственным за действие (в другом смысле. – Д. В.), даже если он не заслуживает за него порицания или похвалы...» [Pereboom 2014b, 2]. Только ответственность в смысле базовой заслуги в случаях индетерминации ставится под сомнение с помощью Аргумента исчезающего агента. Более того, такая ответственность, по мнению Перебума, не может существовать только в *событийно-каузальном либертарианстве (event-causal libertarianism)*.

с) По замыслу Перебума, Аргумент критикует только *событийно-каузальное либертарианство*

Среди современных каузальных либертарианских теорий можно выделить как минимум два типа. Они различаются по сущностям, которые эти теории допускают к участию в каузальных связях. К *событийно-каузальному либертарианству* относятся теории, в которых каузальные связи существуют только между событиями. Согласно этим теориям, действия, анализируемые как события, в которых задействован агент, в качестве вероятностной недетерминирующей причины имеют желания и убеждения, которые также анализируются как события, в которых участвует агент. К сторонникам этого типа либертарианства относятся Кейн, Балагер, Франклин и др. Ко второму типу, *агент-каузальному либертарианству (agent-causal libertarianism)*, относятся теории, в которых каузальные связи могут существовать не

только между событиями, но и между субстанциями. Согласно этим теориям, свобода воли, необходимая для моральной ответственности (в смысле базовой заслуги), обеспечивается каузальной связью между агентами, которые несводимы к событиям, и действиями. При этом сами агенты каузально недетерминированны, то есть обладают свободной волей в либертарианском смысле. К сторонникам этого типа либертарианства относятся О'Коннор, Кларк, Гриффит и проч.

Оба типа либертарианства предполагают необходимость расширенного контроля для ответственности в смысле базовой заслуги. В обоих типах либертарианства такой расширенный контроль обеспечивается индетерминизмом. Но агент-каузальное либертарианство, по мнению Перебума, не может быть мишенью для Аргумента исчезающего агента. Этот аргумент критикует наличие расширенного контроля только в событийно-каузальных теориях либертарианства, и в первую очередь – в специальных случаях, когда силы конфликтующих мотивов агента находятся в равновесии. Аргумент по своему замыслу обнаруживает внутреннее противоречие этого типа либертарианства в данных случаях. Это – противоречие между утверждением необходимости расширенного контроля агента над собственными действиями и фактическим отсутствием такого контроля в индетерминистических ситуациях в связи со случайностью выбора.

Из уточнений, сделанных выше, может сложиться впечатление, что аргумент Перебума имеет узкую направленность:

его вывод относится к одному типу случаев, он опровергает только частный вариант либертарианства и критикует возможность моральной ответственности только в смысле базовой заслуги. Но это не так. Во-первых, случаи равновесных конфликтующих мотивов являются парадигмальными для либертарианцев. Вероятней всего, выводы, относящиеся к этим случаям, могут быть распространены на все остальные случаи вероятностной каузальности, то есть на все теории, основывающие свободу в индетерминизме. Во-вторых, событийно-каузальное либертарианство, которое в этом аргументе критикует Перебум, является, по мнению большинства, самым перспективным вариантом либертарианства. Агент-каузальное либертарианство многим кажется слишком экстравагантным. В-третьих, возможно, что аргумент, использованный в данном случае против событийно-каузального либертарианства, может также быть использован против агент-каузальной теории. Таким образом, Аргумент исчезающего агента представляет проблему для совместимости свободы воли с индетерминизмом в целом. Но, по мнению Перебума, она также несовместима и с детерминизмом.

2. Свобода воли и детерминизм несовместимы: аргумент манипуляций

Перебум, как и Кейн, считает, что детерминизм несов-

местим со свободой воли. Но основания этого убеждения у двух философов отличаются. Кейн, напомним, считает, что моральная ответственность предполагает и наличие альтернативных возможностей, и то, что агент должен обладать «ответственностью в конечном счете», быть единственным источником своих действий. А детерминизм несовместим с двумя этими условиями. Перебум делает вывод о несовместимости детерминизма и моральной ответственности лишь на основании несовместимости детерминизма с условием «ответственности в конечном счете». Он видит несовместимость только «по источнику». Его вывод основывается на *Аргументе манипуляций*³⁵. Он построен на интуиции об отсутствии ответственности и свободы в случае манипуляции. Согласно этой интуиции, если действия агента являются результатом манипулирования, то агент не несет ответственности за действия. Но философ считает, что манипулирование и детерминация сходны между собой. А это значит, что как в случае с манипуляцией, так и в случаях с каузальной детерминацией поведение агента лишено необходимого контроля и, следовательно, агент не несет моральной ответственности.

Аргумент манипуляций Перебума строится на анализе четырех ситуаций. Все четыре ситуации задуманы так, чтобы в них удовлетворялись компатибилистские условия свободы, а интуиции склоняли к суждению об отсутствии моральной

³⁵ Одно из изложений этого аргумента можно найти в книге [Pereboom 2014b, 74].

ответственности. Во всех ситуациях профессор Плам совершает убийство Уайта из корыстных соображений. Перебум начинает исследовать интуиции с умозрительного, но очевидного, по его мнению, случая, где агент является марионеткой в руках ученых. В этой гипотетической ситуации ученые управляют каждым действием агента, и потому агент, кажется, не несет ответственности. Далее Перебум приводит другие случаи манипуляций, все более и более близкие к реальности. Наконец, он доводит примеры до обычного случая, где агентом «манипулируют» не другие агенты, а каузальные связи и физические законы. Философ утверждает, что существенной разницы между всеми этими случаями нет и что экстраполяция интуиции на обыденные примеры должна убедить нас в том, что ответственности нет и в этих, обычных условиях. Итак, первая ситуация – случай, где Плам является марионеткой ученых:

Сценарий 1. Радиоуправляемый агент. Команда нейроученых с помощью радиопередатчиков управляет мозгом и, соответственно, ментальными состояниями Плама. Перед тем как Плам начинает размышлять о действиях, нейроученые нажимают кнопку. Это приводит к тому, что рассуждения Плама принимают крайне эгоистический характер. В результате этих рассуждений Плам решает убить Уайта. Плам не убил бы Уайта, не будь он под управлением нейроученых. Тем не менее эгоистическое решение Плама не противоречит склонностям его характера, так как его характер в це-

лом располагает к эгоистическим соображениям. Желание совершить убийство не противоречит более высокопорядковым желаниям Плама. То есть Плам желает убить Уайта и желает иметь это желание. Желание совершить убийство основано на рациональных мотивах. И будь у него другие рациональные мотивы, он бы воздержался от преступления. Не все рассуждения Плама исключительно эгоистические. Плам также не испытывает «непреодолимых желаний». Нейроученые не внушают ему желаний такого рода [Pereboom 2013a, 424].

По замыслу Перебума, радиоуправляемый агент должен представляться читателям как агент, лишенный моральной ответственности. Ведь его действия – результат манипуляций. При этом манипулируемый Плам по основным параметрам удовлетворяет критериям моральной ответственности, выдвигаемым компатиби-листами. В частности, он удовлетворяет критериям, которые выдвигают Юм, Айер, Франкфурт, Уоллас, Фишер и Равицца. Это должно продемонстрировать, что данные компатибилистские теории не предлагают достаточных условий моральной ответственности. Но могут ли вообще компатибилистские теории предоставить такие критерии? Для ответа на этот вопрос Перебум предлагает рассмотреть сходные, по его мнению, ситуации.

Сценарий 2. Запрограммированный агент. Плам во всех отношениях – обычный человек, кроме одного обстоятельства. В начале его жизни команда нейроученых запрограм-

мировала мозг Плама так, что его рассуждения принимают крайне эгоистический характер. Неизбежным следствием этих рассуждений является решение Плама убить Уайта. Их нейрофизиологическая реализация и история состояний Плама полностью соответствуют Сценарию 1. Эгоистическое решение Плама не противоречит в целом склонностям его характера. Желание совершить убийство соответствует более высокопорядковым желаниям Плама. Желание совершить убийство основано на рациональных мотивах. И будь у него другие рациональные мотивы, он бы воздержался от преступления. Не все рассуждения Плама исключительно эгоистические. Плам также не испытывает «непреодолимых желаний».

Второй Сценарий во многом близок первому. Все выдвинутые ранее компатибилистские критерии свободы воли удовлетворены. А интуиции тем не менее склоняют к признанию отсутствия моральной ответственности. Разве повинен запрограммированный агент в преступлении? Как и в предыдущем случае, разве эта вина не должна быть переложена на злоумышленников-экспериментаторов? Разница между Сценарием 1 и Сценарием 2 – лишь во времени, когда происходит манипуляция. Радиоуправляемым агентом ученые манипулируют в реальном времени. А запрограммированный агент изначально создан так, чтобы совершить преступление. Но эти различия в отношении моральной ответственности, по мнению Пере-бума, ничего не меняют. Как,

впрочем, ничего не должны менять и различия между вторым и третьим Сценариями.

Сценарий 3. Обусловленный агент. Плам – нормальный человек, кроме одной особенности. В детстве он подвергся специальному воспитанию, которое каузально детерминировало его склонность к рассуждениям крайне эгоистического характера. Это воспитание было проведено до того, как Плам стал самостоятельным в своих рассуждениях. Неизбежным следствием эгоистических рассуждений Плама стало его решение убить Уайта. Во всех остальных характеристиках Плам в Сценарии 3 соответствует Пламу в Сценариях 1–2. Плам может руководствоваться моральными доводами, его первопорядковое желание соответствует его второпорядковым желаниям. Эгоистические склонности не противоречат в целом характеру Плама. И у него нет непреодолимых желаний.

Третий Сценарий еще ближе к реальным условиям. Нет ничего удивительного в том, что воспитание в детском возрасте формирует ценности и склонности взрослых. При таком специфическом воспитании ответственность, кажется, переходит от воспитанника к воспитателям. То есть и в третьем Сценарии, несмотря на удовлетворение компатибилистских критериев свободы воли, ответственность у агента отсутствует. Перебум считает, что такие интуиции должны распространиться и на четвертый, заключительный случай:

Сценарий 4. Детерминированный агент. Все, что проис-

ходит в мире, каузально детерминировано событиями прошлого и законами природы. Плам – обычный человек, родившийся и воспитанный в обычных условиях. Его рассуждения часто (но не всегда) носят крайне эгоистический характер. Его решение убить Уайта происходит из склонностей его характера и формируется в ходе рациональных рассуждений. Это его желание полностью соответствует его второпорядковому желанию иметь такое первопорядковое желание. Нейрофизиологическая реализация и история состояний Плама полностью соответствуют Сценариям 1–3. Плам способен в целом регулировать свое поведение моральными принципами и не действует из непреодолимых желаний.

Если рассуждать последовательно, то и в четвертом случае моральная ответственность должна отсутствовать. Ведь этот случай во многом аналогичен предыдущим. Сценарии подобраны таким образом, что каждый последующий лишь незначительно отличается от предыдущего, притом несущественным для признания наличия ответственности образом. Перебум сначала представляет оригинальный контрпример к компатибилистским критериям свободы воли, а затем обобщает его до случая со стандартными, нормальными условиями. В процессе этого обобщения, выстраивая все новые и новые примеры, Перебум выявляет ключевую причину отсутствия моральной ответственности. Она заключается, по его мнению, в каузальной детерминации действий агента. Каузальная детерминация, действующая во всех че-

тырех случаях, снимает ответственность с агента, так как она находится вне зоны его контроля. Агент не может контролировать ни предшествующие события, ни законы природы, и именно поэтому не несет ответственности за производимые действия. Из этого следует, что не только учтенные компатибилистские теории не могут представить достаточных условий свободы воли и ответственности, но и любые компатибилистские теории в принципе не могут представить таких критериев. Свобода воли и моральная ответственность несовместимы с каузальной детерминацией. Компатибилизм ложен.

Аргумент Перебума можно представить в более формальном виде. Аббревиатурой «ККС» обозначим компатибилистские критерии структуры свободного агента. Под ККС будем подразумевать все необходимые условия, которым, в соответствии с компатибилистскими теориями, должно отвечать действие, чтобы его можно было признать свободным. В число этих критериев входят соответствие действия склонностям характера, отсутствие непреодолимых желаний, соответствие первопорядковых желаний второпорядковым, согласованность желаний с рациональными мотивами агента. С учетом этих критериев, аргумент будет выглядеть так:

1. Если агентом S успешно манипулируют, чтобы он сделал A , тогда S делает A не по своей воле, и потому он не несет моральную ответственность за действие A .

$(M \rightarrow \neg R)$

2. Манипулируемый агент, выполняющий действие A с удовлетворением ККС, не отличается в релевантном смысле от обычного каузально детерминированного агента. $(M = D)$

3. Следовательно, обычный агент, детерминированный и удовлетворяющий ККС, не действует по своей воле, и потому не несет моральной ответственности. $(D \rightarrow R)$

Первая посылка кажется интуитивно верной. К тому же она иллюстрируется первым Сценарием. Вторая посылка опирается на обобщение от Сценариев 1–3 к Сценарию 4. При истинности посылок, вывод тоже кажется верным. Аргумент выглядит изобретательным и правдоподобным. Означает ли это, что свобода воли и ответственность несовместимы с детерминизмом? Более того, в совокупности с Аргументом исчезающего агента это может означать, что свобода воли и ответственность вообще не совместимы ни с детерминизмом, ни с индетерминизмом.

Я не согласен с этим выводом и собираюсь доказать, что свобода воли и ответственность совместимы как с детерминизмом, так и с индетерминизмом. Эта позиция называется *суперкомпатибилизмом*. Часть работы уже проделана – были представлены соображения против Аргумента последствий в отношении «способности поступить иначе». Я также показал способ защитить «примеры в стиле Франкфурта». Эти примеры доказывают отсутствие необходимости доступа к

альтернативным возможностям для свободы и ответственности. Теперь, чтобы отстаивать позицию суперкомпатибилизма, осталось показать, что Аргумент манипуляций, направленный против совместимости детерминизма и свободы, и Аргумент исчезающего агента, обращенный против совместимости индетерминизма и свободы, несостоятельны. Это я намерен сделать в заключительной, четвертой части этой главы.

IV. Обоснование суперкомпатибилизма

1. Возражение на Аргумент манипуляций

Чтобы проверить надежность Аргумента манипуляций, нужно разъяснить некоторые моменты. Во-первых, нужно убедиться в правомерности обобщения всех случаев с манипуляциями и обыденного случая. Действительно ли Сценарии в релевантном смысле сходны между собой и в особенности с четвертым Сценарием? В чем это сходство, а в чем – отличия? Во-вторых, нужно убедиться, что случаи с манипуляциями действительно снимают моральную ответственность с агента. И в чем, наконец, основание интуиций о моральной ответственности? Для того чтобы все это выяснить, нужно действовать как экспериментаторы в лаборатории. Нужно провести несколько преобразований со Сценариями Перебума, поочередно внося в них изменения, и посмотреть, что получится. Определим для начала, какие параметры можно варьировать.

В сценариях Аргумента манипуляций есть несколько ключевых параметров. На первый взгляд это: (1) *тип манипуляторов*; (2) *время манипуляции*; (3) *компатибилистские критерии свободы агента*; (4) *характер действия*; (5) *последо-*

вательность сценариев; (б) тип детерминации. Изменяя эти параметры, можно попытаться выяснить, что будет происходить с интуицией о моральной ответственности. Представьте, что каждый из этих параметров – тумблер на панели моральной ответственности. Мастер-выключатель выключает все тумблеры. При выключении мастер-выключателя электрическая лампочка гаснет – появляется интуиция, что моральная ответственность исчезает. Следует выяснить, какой из обычных тумблеров отвечает за отключение лампочки, какой из параметров отвечает за интуицию об отсутствии моральной ответственности. Перебум считает, что этим тумблером является тумблер каузального детерминизма. Проверим это, включая и выключая тумблеры по очередности.

В Сценариях с первого по третий действия Плама контролируют *другие агенты (1)*, а в четвертом их вовсе нет. В первом Сценарии Пламом манипулируют нейроученые, которые с помощью радиосигналов управляют его мозгом. Во втором – это программисты мозга Плама, в третьем – воспитатели, обуславливающие его поведение. В четвертом манипуляторов нет, и контроль, кажется, остается за законами природы. Можно предположить, что именно наличие других управляющих агентов освобождает Плама от ответственности. Другие агенты, как магнит, перетягивают моральную ответственность в свою сторону. И это и есть отвечающий за интуицию тумблер. Если это так, то аргумент Перебума обречен на неудачу с самого начала, ведь даже если ответствен-

ность действительно отсутствует в первых трех случаях, ее обобщение на четвертый несостоятельно.

Проверим это: уберем агентов из первых трех Сценариев, выключим тумблер «манипуляторы» и... увидим, что радикальных изменений не происходит. Если вместо ученых поставить случайную работу машины, не имеющей специальной цели, или случайное воздействие электромагнитного поля, интуиции сильно не изменятся. По мнению Перебума, эффект, отсутствие моральной ответственности, точно останется. Плам совершит убийство, но моральной ответственности за него не должен будет нести. В этом пункте можно согласиться с Перебумом. Интуиции о наличии моральной ответственности, скорее всего, не возникнут. Хотя в этом случае интуиции могут быть уже не так однозначны. В любом случае очевидно, что причина не в том, что Пламом управляют другие агенты, точнее, по крайней мере, не только в этом, но и, по-видимому, в чем-то другом.

Другой параметр – *время манипуляции (2)*. В первом случае Плам управляется в режиме реального времени, то есть момент манипуляции непосредственно предшествует действию. В третьем Сценарии манипуляция осуществляется в детстве Плама, за много лет до совершения убийства. Во втором Сценарии воздействие происходит еще раньше, до рождения Плама. Четвертый Сценарий предполагает «манипуляцию» в бесконечно далеком прошлом: в момент, когда законы природы были сформированы. То есть манипуляция

может отодвигаться на сколь угодно длительное время от действия. Но имеет ли значение время манипуляции в детерминированном мире? Представим, что манипулятором в первом Сценарии является бог, который создает мир, в котором через тысячелетия будет рожден и воспитан Плам. Если при этом все каузальные связи сохранятся и история развития Плама будет совпадать с историей в первом Сценарии, время манипуляции не будет, кажется, иметь значения. Плам все равно не будет нести ответственность за действия, предопределенные в далеком прошлом. Выходит, тумблер «время» не имеет принципиального значения. Хотя в этом случае у меня интуиция уже не так однозначна. Как бы то ни было, очевидно, что причина не в том, что Пламом управляют в режиме реального времени, точнее, по крайней мере, не только в этом, но и в чем-то другом.

Возможно, тогда дело – в *конкретных компатибилистских критериях* (3) свободного ответственного действия? Попробуем варьировать этот параметр. Перебум использует критерии свободного ответственного действия, которые фигурируют в наиболее развитых компатибилистских теориях. Более того, он использует их в сочетании. В сумме они могли бы обеспечить необходимые и достаточные условия свободы действия. Но, кажется, они в лучшем случае обеспечивают только необходимые условия. Даже в самом первом Сценарии эти условия могут быть полностью удовлетворены, но при этом моральной ответственности нет. Может, тогда сто-

ит пополнить или изменить их список? По мнению Перебума и даже некоторых компатибилистов, пополнение ККС не изменит интуиций принципиально. Почему? Допустим, существует «лучшая компатибилистская теория», в которой учтено условие X , являющееся лучшим кандидатом на достаточное условие свободного ответственного действия. Все равно кажется возможным модифицировать все сценарии Аргумента манипуляций так, чтобы они учитывали это условие X , но чтобы при этом агент в них все равно был лишен ответственности. Просто в таком случае манипуляция будет более сложная. С этим тоже в принципе можно согласиться. Тумблер ККС не выключает лампочку.

Следующим идет тумблер «тип поступка» (4). Философия свободы воли изобилует примерами тяжелых преступлений. Первое место, пожалуй, занимают случаи с убийствами. Поэтому порой философские произведения выглядят как гибриды интеллектуальной прозы и жестоких второсортных детективов. Объяснение у этого простое. Философам нужны примеры, которые вызывали бы однозначную интуитивную реакцию. В этом смысле убийства подходят. Но что произойдет, если убийство Уайта заменить другим действием – например, действием, заслуживающим одобрения и поощрения? Предлагаю новый сценарий с другим типом действия.

Сценарий 5. Радиоуправляемый герой. Команда нейроученых с помощью радиопередатчиков управляет ментальными

состояниями Плама. Плам видит, что из окон соседнего дома идет дым. От соседей он узнает, что в горящем доме – четырехлетний ребенок. Перед тем как Плам начинает думать о том, как поступить, нейрочувствительные нажимают кнопку. Это приводит к тому, что рассуждения Плама принимают альтруистичный характер. Плам решает спасти ребенка. Он разбивает окно и, прикрываясь от огня и дыма курткой, попадает в дом. В доме горят вещи, воздух раскален и пропитан едким дымом. Плам обжигает руки и лицо. Но это его не останавливает. В спальне он находит маленького испуганного мальчика, сидящего на полу. Плам берет его на руки и выносит на улицу. Спасенный мальчик обнимает Плама, родители благодарят его и восхищены мужественным поступком. Плам отказывается принимать благодарности и скромно прощается с новыми знакомыми. Он чуть не погиб сам, спасая ребенка, но, если бы аналогичный случай произошел еще раз, Плам бы поступил точно так же.

Во всех остальных аспектах этот сценарий похож на Сценарий 1 с радиоуправляемым агентом. Поступок Плама не противоречит в целом склонностям его характера. Его желание не противоречит его второпорядковым желаниям и основано на рациональных мотивах – спасти ребенка. Несет ли альтруистичный Плам моральную ответственность? Заслуживает ли он похвалы, благодарности?

Кому-то, возможно, покажется, что нет. Но у меня – в большей степени другие интуиции. Мне кажется, что с этой

модификацией Сценариев обнаруживается любопытная особенность. При одном и том же механизме манипуляций оценки будут разные в отношении Сценариев с положительными и отрицательными поступками. Да, характер Плама стал более альтруистичным в результате манипуляции, но ведь он все-таки стал таковым, и Плам совершил хороший поступок, имея добрые мотивы. Если Плама вообще можно считать агентом, и этот поступок можно считать его поступком, он, с моей точки зрения, заслужил уважения и благодарности.

Подобные же интуиции разделяют и американские философы-компатибилисты С. Вульф и Д. Нелкин. Это – странная ситуация. Выходит, изменение характера поступка даже в случаях с манипуляциями влияет на признание наличия моральной ответственности? Но к какому выводу это нас толкает? К убеждению, что иммунитетом в отношении Аргумента манипуляций обладают только положительные поступки и что только моральная ответственность за положительные действия совместима с каузальным детерминизмом? Или, вообще, что моральная ответственность оправдана только в случае с положительными поступками? Не странно ли это?

У Перебума есть на это ответ. «Я согласен, что интуитивно кажется более уместным хвалить Плама в этом примере [с положительным поступком], чем в исходном Сценарии... Но для меня не столь очевидно, что эта похвала и мораль-

ное одобрение являются признаками моральной ответственности в смысле базовой заслуги...» [Pereboom 2014b, 102]. Людям в принципе свойственно хвалить других за врожденный интеллект, внешнюю привлекательность, таланты. Хотя эти качества – вовсе не заслуга агентов. Похвала в данном случае, по мнению философа, – это выражение восторга, радости, позитивного отношения, а не признак моральной ответственности. Возможно, и в случае с модифицированным Сценарием реакция на поступок Плама – это просто выражение радости в отношении спасенного ребенка. Другими словами, Перебум считает, что этот тумблер не является искомым, что он не влияет на лампочку моральной ответственности. Это – ложноположительный результат. Допустим. Тогда переключим следующий тумблер.

Поэкспериментируем с еще одним параметром: (5) *последовательностью Сценариев*. Перебум утверждает, что все ситуации аналогичны между собой в отношении моральной ответственности. Следовательно, их можно чередовать в любом порядке. Тогда разместим их в обратной последовательности. Начнем анализ со случая самого обычного – *Детерминированного агента*. Какие интуиции возникают: несет ли агент моральную ответственность? Многие могут решить, что в этом случае моральная ответственность существует. Такое отношение к данному случаю исходит из нормальной практики человеческих отношений. Но даже более скептическая позиция приемлема. Допустимо считать, что в си-

туации с Детерминированным агентом неясно, является ли этот агент свободным или нет и, соответственно, несет ли он ответственность. Теперь развернем второй случай, случай с *Обусловленным агентом*. Если тактика обобщения работает, и Сценарии не отличаются в релевантном смысле друг от друга, случай с Обусловленным агентом должен опять привести к агностическому суждению. Но если продолжать дальше, то и в отношении запрограммированного и даже манипулируемого агента будет сохраняться агностическая интуиция. Такую инверсию Аргумента манипуляций предлагает другой компатибилист, М. Маккенна. Его вывод нейтрален: «Я пытаюсь продемонстрировать, что из Сценариев... не следует, что Плам не свободен и не несет моральной ответственности» [McKenna 2008, 153]. Из одного только детерминизма вывести отсутствие свободы воли и ответственности нельзя. Но если тактика обобщения верна, тогда агностический вывод достоверен.

Однако в чем достоинство агностической позиции? Предположим, из Аргумента следует, что моральная ответственность может быть, а может и не быть совместимой с детерминизмом. Тогда что? Маккенна считает, что уже это – частичная победа. Ведь задача Аргумента манипуляции – именно в том, чтобы представить основания для жесткого инкомпатибилизма. А если вывод из аргумента агностический, то этих оснований нет. Аргумент перестает выполнять необходимую роль. Естественно, Перебум с этим не может согласиться.

По его мнению, результат анализа Аргумента и его модифицированных инвертированных вариантов зависит от исходных установок. Так, некоторые начинают рассуждения с позиции убежденного агностика. Сторонники этой позиции считают, что в отношении моральной ответственности неясно, совместима она с детерминизмом или нет. Но убежденные агностики не видят возможности эту позицию дальше прояснить. Для них на этом вопрос уже закрыт. Можно также начинать рассуждения с позиции убежденного компатибилиста или убежденного инкомпатибилиста. Но и эти позиции не предполагают продуктивной дискуссии, заведомо предрешая итоговый вывод. Более продуктивны открытые исследовательские позиции, ищущие согласование с новыми доводами и фактами. Перебум считает, что агностический вывод из аргумента Маккенны следует только из позиции убежденного компатибилиста, закрытого для принятия новых фактов. А новые факты здесь имеют значение. Эти факты свидетельствуют о том, что основания для поведения агентов лежат за пределами самих агентов, даже за пределами их жизненных историй. В связи с этим Перебум неоднократно ссылается на Спинозу и на его объяснение иллюзии свободы: люди в опыте находят подтверждение своей уверенности в свободе, так как они осведомлены о своих желаниях и действиях, но не знают о причинах, их вызывающих. Именно эти новые факты, свидетельствующие о детерминирующих обстоятельствах, должны склонить чаши весов

в сторону инкомпатибилизма. И как раз это должен продемонстрировать Аргумент манипуляций.

Маккенна апеллирует к компатибилистическим интуициям в обыденной жизни. И предлагает больше обращать внимание на агентские способности Плама в Сценариях. Перебум настаивает на том, что нужно сфокусироваться на внешних обстоятельствах: на манипуляторах, программистах поведения и цепочке причин. И действительно, смена фокуса влияет на восприятие Сценариев. Интуиции, как зрительные образы в случае с кубом Неккера, изменяются в зависимости от исходных установок и акцентов. Кажется, что Аргумент манипуляций заводит в тупик. Как разрешить эту ситуацию? Пробуем переключить последний, шестой тумблер.

Это *тип детерминации (6)*. Напомню, во всех Сценариях Аргумента причины с необходимостью определяют результат, то есть агенту доступно только одно будущее. Агент и его манипуляторы существуют в мире с каузальным детерминизмом. Что произойдет, если в Сценариях ввести индетерминированные события? Появится ли интуиция о моральной ответственности? Как ни удивительно, последствий в отношении моральной ответственности, кажется, тоже не будет. Даже если манипуляторы Плама будут определять его действия лишь с некоторой вероятностью, интуиции в отношении его ответственности не будут поколеблены. Пусть в одном случае манипуляторы управляют его действиями со стопроцентным успехом, в другом – только с некоторой ве-

роятностью, – для Плама это ничего не меняет.

Это – любопытная ситуация. Она, казалось бы, подрывает аргументацию Перебума. Ведь в конечном итоге с помощью Аргумента манипуляций он критикует компатибилизм. К тому же философ, как я уже говорил, сам утверждает, что причиной отсутствия моральной ответственности как в Сценариях 1–3, так и в четвертом, стандартном случае, является каузальная детерминация. Но вот мы убираем каузальную детерминацию, вносим индетерминизм, и ничего не меняется. Как такое может быть? На это, однако, у Перебума есть ловкий и вполне убедительный ответ. Каузальная детерминация в данных конкретных случаях действительно является причиной отсутствия моральной ответственности. Но это – только в конкретных случаях. Есть еще и более глубокая, общая причина, которая подрывает ответственность. Это – присутствие в каузальной истории событий, над которыми у агента нет контроля. Эта более общая причина объясняет отсутствие моральной ответственности как в случаях с каузальной детерминацией, так и в случаях с индетерминированными событиями.

Для того чтобы прояснить соотношение между двумя этими случаями, Перебум приводит такую аналогию. Предположим, предлагает он, где-то прорвало плотину. Инженера спрашивают, в чем причина. Инженер объясняет, что плотину прорвало, так как резервуар был наполнен более чем миллионом тонн воды. На это его начальник возражает: вода

здесь ни при чем. Представь, что резервуар был бы наполнен более чем миллионом тонн нефти. Ведь тогда бы произошло то же самое? Значит, дело не в воде. Судя по всему, начальник в этом примере не понимает, что здесь нет противоречия. В конкретном случае плотина прорвалась от переполнения резервуара водой. Но она действительно прорвалась бы, даже если резервуар переполнился бы нефтью. Ведь есть еще более глубокая, общая причина прорыва: превышение силы давления на стену плотины. Прорыв плотины, как и отсутствие моральной ответственности, имеет причины разных уровней общности и глубины. В описанных ситуациях, отмечает Перебум, причиной является наличие каузальной детерминации. В других ситуациях ею мог быть индетерминизм, но в любом случае общей причиной является отсутствие необходимого контроля.

Думаю, с этим вряд ли можно согласиться. И тумблер, отвечающий за исчезновение интуиции о моральной ответственности в Сценариях, – это не (6) «*каузальная детерминация*», а (7) «*личность*» Плама. Интуиция о моральной ответственности возникает тогда, когда мы представляем Плама единой целостной личностью, проживающей последовательную, достаточно полную историю, и исчезает, когда этой целостности и последовательности нет. На мысль об этом наводит уже проделанная вариация типа поступка.

Напомню, Перебум говорит, что благодарность в случае альтруистического поступка – это не свидетельство интуи-

ции о наличии моральной ответственности, а просто реакция на стечение обстоятельств. Но почему нужно так считать? Мне кажется, что это именно реакция в отношении поступка Плама. И она справедлива. Одно дело – выражать радость в отношении каких-то обстоятельств, но совсем другое – реагировать, в частности с благодарностью, на поступки агента. Реакция в отношении Плама положительная, и эта реакция направлена на него, на его поступок, а не на обстоятельства ситуации. И это – показатель наличия моральной ответственности в смысле «заслуги». Здесь я согласен с Вульф и Нелкин.

Но мне кажется неверной их интерпретация причины. Эти философы предполагают, что моральная ответственность асимметрична. То есть что каузальная детерминация несовместима с отрицательными реакциями, но совместима с положительными. Именно поэтому моральная ответственность в случаях с отрицательными действиями отсутствует, а в случаях с положительными – присутствует. По моему мнению, оснований для утверждения асимметрии недостаточно. Есть другое, более простое объяснение. Когда манипулируемый Плам поступает альтруистично, он переживает последствия собственных действий. Бросаясь спасать ребенка, он оказывается в горящем помещении. Он рискует задохнуться, сгореть, получить ранение, совершая альтруистичный поступок. Решение согласовано с его желаниями, его характером (пусть даже «новым»), жизненной историей и последствия-

ми. И потому он заслуживает похвалы и благодарности. Он – целостная личность, совершающая поступки и испытывающая их последствия.

Проблема с Аргументом манипуляций в том, что в Сценариях Перебума все заканчивается убийством, которое совершает Плам. Но для точной установки интуиции не менее важно, что происходит потом: как реализуются его цели, как он переживает результаты своих действий. Что за личность Плам после поступка? Остается ли он той же личностью или изменяется? Предполагаю, что интуиции в отношении ответственности Плама-убийцы могут пошатнуться, если подробнее описать его дальнейшее поведение и обстоятельства жизни.

Продолжение к Сценарию 1. Закоренелый радиоуправляемый преступник. В результате сдвига системы ценностей под воздействием манипуляции Плам совершает преступление. Он сначала похищает Уайта, пытается его, заставляя подписать бумаги, а потом жестоко убивает. Тело Уайта Плам сбрасывает в реку. Покончив с уликами, Плам оформляет имущество Уайта на свое имя. Он переезжает в усадьбу Уайта и начинает наслаждаться жизнью. Он снимает деньги с банковского счета Уайта, тратит их на наркотики и азартные игры. А когда кто-то из родственников Уайта начинает догадываться об убийстве, Плам совершает еще одно преступление. Ему не нужны свидетели и конкуренты. Это происходит, конечно, уже без каких-то дополнительных манипу-

ляций со стороны нейроученых. Они ведь еще до первого убийства сдвинули его шкалу ценностей в сторону крайней эгоистичности. Дальше все разворачивается как естественное продолжение.

Неужели и теперь у нас должно сохраниться ощущение невинности Плама? Да, Плам в некотором смысле – жертва обстоятельств. Но он уже стал отвратительной личностью и продолжает ею оставаться. Его поступки сообразны его личным качествам, его характеру и ценностям. Их непосредственно относят к Пламу. Разве по справедливости он не должен отвечать за преступления? Другое дело, если бы после убийства Уайта Плам оказался морально раздавленным. Если бы он винил себя, отчаянно пытался разобраться, как все это произошло, и отправился бы с признанием в полицию. В таком случае, зная об особых обстоятельствах, мы оценили бы поступок как результат манипуляции, как случайность, как ситуацию, в которой Плам был «сам не свой». Тогда бы в исторической целостности Плама был перебой. И тогда бы мы сняли с Плама моральную ответственность за содеянное.

В общем, описания ситуаций Перебумом в целом не дают достаточно информации о том, какой личностью после поступка остается Плам. Остается ли он после манипуляции все таким же эгоистичным, беспринципным, жестоким, корыстным? Или после манипуляции он резко меняется, горько сожалеет, корит себя за внезапный порыв жестокости?

Мы также не знаем, как действуют дальше нейроученые: продолжают ли они свои манипуляции или довольствуются первым изменением. Если с их стороны больше воздействий нет, и Плам сохраняет целостность личности и когерентность личностной истории, моральная ответственность, по моему мнению, сохраняется. А если воздействия продолжаются, или Плам резко меняется сам, возвращаясь к исходной системе ценностей, моральной ответственности быть не должно.

Таким образом, из этих рассуждений можно сделать следующий вывод: Аргумент манипуляций внушает интуиции об отсутствии моральной ответственности не из-за детерминизма, а из-за информационных лакун, допущенных в мысленном эксперименте. Именно эти лакуны не позволяют судить о Пламе как о самостоятельной ответственной личности. А ведь для моральной ответственности нужен не только проступок и злой умысел, но и целостная аморальная личность, которой можно приписать поступок и когерентную личностную историю. Детерминизм не противоречит существованию таких целостных личностей. А значит, Аргумент манипуляций не работает против компатибилизма.

2. Возражение на аргумент исчезающего агента

С моей точки зрения, не обладает убедительностью и Ар-

гумент исчезающего агента. Проблемы этого аргумента обнаруживают оппоненты Перебума: А. Мили и К. Франклин. Их возражения мне кажутся очень полезными, хотя и не решающими. Ниже я предложу свой вариант критики, основанный на одной из возможных интерпретаций понятия личности. Этот вариант мне представляется наиболее перспективным.

А. А. Мили

а) Аргумент против событийно-каузального либертарианства действует также против агент-каузального либертарианства

Перебум утверждает, что Аргумент исчезающего агента направлен только против событийно-каузального либертарианства, но не может опровергнуть агент-каузальную теорию [Pereboom 2014b, 50]. В первой теории требуется наличие расширенного контроля для моральной ответственности, но Аргумент показывает, что в рамках этой теории такого контроля у агента нет. Агент не устанавливает, какие решения он примет в ситуации индетерминизма. Подобной проблемы, по мнению Перебума, нет во второй теории. В ней агент даже в ситуации индетерминизма может устанавливать, какое решение будет принято. Он обладает расширенным контролем. Этот контроль обеспечивается субстанциональной каузальностью агента. Но действительно ли агент-

каузальная теория обладает иммунитетом по отношению к Аргументу исчезающего агента? Действительно ли субстанциональная каузальность спасает эту теорию?

Один из давних оппонентов Перебума, А. Мили, оспаривает это утверждение. Философ считает, что, если проблема случайности выбора есть в событийно-каузальном либертарианстве, она есть и в агент-каузальном либертарианстве [Mele 2017]. Мили предлагает свой анализ ситуации спорного решения уже в системе агент-каузального либертарианства. Предположим, пишет Мили, агент каузально определяет решение. Но в ситуации спорного выбора (например, в описанных выше Случаях Кейна «Ограбление» или «Выбор Ральфа»), где каждое из альтернативных решений может быть принято с вероятностью 0,5, в половине случаев агент сделает выбор в пользу альтернативы *A*, а в другой половине случаев – в пользу альтернативы *B*. На чем будет основываться различие в выборе?

Ментальные состояния агента по определению одинаковы, мотивы – тоже. Ничто не отличается в ситуациях с разными исходами. Что тогда определяет, какой выбор сделает агент? Сам агент? Но чтобы выбор был различным, различия должны быть в самом агенте, а их, по условиям, нет. Кажется, ничто не может объяснить разницу в исходах. Но если нет ничего, что могло бы объяснить эту разницу, выбор представляется случайным. Следовательно, если Аргумент исчезающего агента опасен для событийно-каузально-

го либертарианства, он опасен и для агент-каузальной теории. Агент-каузальное либертарианство, видимо, тоже не дает расширенного контроля. Но есть и другая интерпретация: мы просто не очень хорошо понимаем теорию агент-каузального либертарианства – в ней есть скрытые аксиомы.

б) Агент-каузальное либертарианство предполагает полный контроль над выбором, но полный контроль над индетерминистическим событием невозможен

В целом кажется, что агент-каузальное либертарианство – очень смутная теория. Об этом говорят даже сами либертарианцы, такие как Р. Кейн. Тем не менее в качестве гипотезы можно представить, как это готов сделать Перебум, что в агент-каузальном либертарианстве агентная каузальность не допускает случайности. Тогда, видимо, агентная каузальность должна давать агенту полный контроль над исходом решений. Для того чтобы решение не было случайным, чтобы агент был свободен в той степени, что необходима для моральной ответственности, у него должен быть полный контроль над исходом выбора. Но не слишком ли это высокие требования? Неужели мы всегда ожидаем, что агент обладает полным контролем, когда приписываем ему ответственность?

Предположим, профессиональный баскетболист имеет 90 %-й шанс забросить штрафной. У него есть большой контроль над тем, чтобы забросить мяч, продолжает Мили.

«Утверждение, что, несмотря на это, он не может свободно забросить любой из штрафных мячей потому, что у него нет полного контроля... устанавливает слишком высокую планку...» [Mele 2017, 566]. Возможно, в случае с либертарианским агентом полный контроль нужен не над исходом выбора, а над попыткой сделать выбор? Тогда представим другой сценарий: баскетболист имеет 90 %-й шанс *сделать попытку* забросить мяч и 10 %-й шанс психологического срыва – в его голове установлен индетерминистический генератор случайных чисел с соответствующими вероятностями исхода. Вот он забрасывает мяч. Можно ли сказать, что он забросил этот мяч совершенно случайно? Мили считает, что и в этом случае требование полного контроля – чрезвычайно строгое требование. К тому же философ видит проблему в представлении о полном контроле индетерминистического выбора.

Полный контроль над выбором *A*, видимо (по аналогии с полным контролем над действием), предполагает, что у агента есть намерение сделать выбор *A*, что он пытается сделать выбор *A* и не может не преуспеть, в результате делая выбор *A*. Но в действительности агент почти никогда не имеет «намерения сделать выбор *A*» и не «пытается сделать выбор *A*». Агент сразу делает выбор *A*. Это более экономичный вариант мышления. Было бы странно, если бы наличие моральной ответственности предполагало бы такой неэффективный способ мышления. Кроме того, полный контроль у агента должен быть над выбором, который при этом должен

быть индетерминистическим, то есть который может быть не сделан. В этой ситуации невозможно определить, что такое полный контроль, или, скорее, он просто невозможен.

В. К. Франклин: Усовершенствованный вариант событийно-каузального либертарианства позволяет устанавливать выбор

Аргумент исчезающего агента обнаруживает потенциальные проблемы в событийно-каузальном либертарианстве, – соглашается другой комментатор Перебума, К. Франклин. Этот философ тоже считает, что теории многих современных либертарианцев, в частности М. Балагера и Р. Кейна, оставляют выбор агента в условиях индетерминизма случайным. Но он уверен, что эти теории можно усовершенствовать. «У сторонников событийно-каузального либертарианства не получится создать удовлетворительной теории свободы воли, просто наложив поверх каузальной теории действия требование индетерминизма. <Редуктивный анализ свободы воли> они должны дополнить редуктивным анализом способности *самоопределения*» [Franklin 2014, 418]. Это дополнение необходимо, чтобы решить проблему с исчезновением ответственности и свободы. Под *способностью самоопределения* Франклин понимает возможность агента устанавливать или разрешать какой-то спор внутри себя.

Способность самоопределения позволяет агенту управлять своими мотивами и делать выбор в пользу какой-то одной альтернативы действий. Эту способность Франклин (вслед за другими теоретиками свободы воли, Уотсоном, Шумейкером и др.) считает необходимым условием свободы агента. К чему и каким образом можно ее редуцировать?

Поскольку концепция самоопределяющегося агента является концепцией функциональной, Франклин считает необходимым применение функциональной редукции. Функциональная редукция в данном случае заключается в сведении активности самоопределяющегося агента к состояниям и событиям, которые играют его функциональную роль, а также могут быть идентифицированы с агентом. Чтобы агент мог быть идентифицирован с состояниями и событиями, они должны выражать авторитет агента, «и поскольку у этих состояний есть авторитет агента, выполнение этими состояниями функциональной роли самоопределяющегося агента может считаться выполнением им своей функциональной роли» [Franklin 2014, 420].

По мнению Франклина, агент в большей степени может быть идентифицирован со своей рациональностью. Исходя из этого, состоянием, которое может быть идентифицировано с агентом в конкретный момент, Франклин считает состояние «желания действовать в соответствии со своими самыми здравыми соображениями» (*desire to act for what the agent takes to be his strongest reasons*). Это желание свойственно

только самоопределяющимся агентам – более простые агенты, животные или младенцы, не обладают такими желаниями. И соответственно, оно не может быть отчуждено от самоопределяющегося агента без того, чтобы он перестал быть таковым.

Впрочем, даже если усомниться, что именно это состояние является сущностным для агента, можно предположить, что другие сходные состояния играют функциональную роль агента. Например, таким состоянием может быть желание действовать в соответствии с самыми сильными чувствами или заботами. Франклин предполагает, что состояния, выражающие самоопределяющегося агента, могут иметь *множественную реализацию*, то есть у разных агентов роль функции самоопределения могут играть разные состояния. В любом случае наличие такого состояния в дополнение к основным мотивам важным образом меняет обстоятельства выбора. Это можно обнаружить, если применить описанную выше концептуальную схему, например, к ситуации с выбором Ральфа.

Представим, что в результате размышления над собственными мотивами Ральф приходит к выводу, что его амбиции стать звездой в Нью-Йорке являются самыми здравыми соображениями. Ральф решает действовать в соответствии со своими самыми здравыми соображениями – он совершает самоопределяющее действие. В таком случае мотивационная сила амбиций Ральфа дополняется мотивацион-

ной силой действовать в соответствии с самыми здравыми соображениями. Это – новый каузальный фактор, определяющий действия агента. Это желание можно идентифицировать с самим агентом, оно представляет самого агента. В нем агент реализует каузальную роль в добавление к каузальной роли своих желаний. Это каузальное влияние устанавливает, в дополнение к базовым мотивам, выбор. Соответственно, по мнению Франклина, этот выбор и последующие действия не являются случайными при общем сохранении индетерминизма в определении исхода выбора. Таким образом, как считает философ, недостаток, выявленный Перебумом, устранен, а Аргумент исчезающего агента опровергнут.

С. Личность как пердурантная сущность устанавливает выбор

Частично мое мнение совпадает с мнением критиков Перебума. Так, я согласен с Мили в том, что индетерминистические обстоятельства в интерпретации Перебума приводят к дилемме: либо в индетерминистических ситуациях для агентной каузальности тоже возникает случайность выбора, либо агентная каузальность внутренне противоречива. С одной стороны, если не допускать наличия полного контроля у субстанционального агента в агент-каузальном либертарианстве, выбор в спорных ситуациях все равно остается случайным. Добавление агента в качестве вспомогательного ка-

узального фактора не позволяет объяснить, что обуславливает именно этот, а не другой выбор. С другой стороны, в принципе полный контроль задает слишком высокую планку для условий моральной ответственности. А в условиях индетерминизма полный контроль вообще, кажется, невозможен.

Я также отчасти согласен с Франклином. Существует возможность усовершенствовать либертарианскую позицию так, чтобы обойти Аргумент исчезающего агента. Даже если либертарианские теории Кейна и Балагера уязвимы для аргумента Перебума, можно представить улучшенную альтернативу. Можно усовершенствовать условия даже индетерминистических ситуаций так, чтобы агент действительно устанавливал выбор и совершал ответственные действия. Для этого необходимо ввести редуктивный анализ способности самоопределения, как предлагает Франклин. В этом я тоже согласен с ним. Однако сама редукция, проведенная Франклином, кажется мне неудачной. По моему мнению, она контринтуитивна и не решает поставленные автором задачи.

Франклин вводит дополнительный каузальный фактор, *метажелание*, которое, по его мнению, в ситуациях самоопределения должно представлять агента. Но это все равно не исключает различных исходов, ведь даже с этим фактором, учитывая индетерминистические обстоятельства, возможно принятие как одного, так и другого решения. При этом мы не можем объяснить, почему агент делает такой выбор, а не иной, а без этого объяснения выбор кажется случай-

ным. И следовательно, агент не устанавливает, какой выбор будет сделан. Но что тогда необходимо, чтобы избежать случайности? Как нужно дополнить описание индетерминистической ситуации, чтобы агент устанавливал выбор? По моему мнению, критерием способности устанавливать выбор является наличие контрастного объяснения, отсылающего к самому агенту.

Контрастное объяснение – это такое описание, которое объясняет, почему в результате выбора случилось именно такое, а не другое событие. То есть отсылка к агенту и его свойствам должна объяснять, почему был сделан именно этот, а не другой выбор. Допустим, мы могли бы проиграть ситуацию с Ральфом бесконечное число раз. Удовлетворительный анализ агентного контроля должен представить объяснение, почему выбор Ральфа отличался в 50 % случаев. С моей точки зрения, ни агент-каузальное либертарианство, ни анализ Франклина в данном случае не позволяют представить контрастное объяснение. Поэтому эти теории все равно не решают главной потенциальной проблемы недетерминированного выбора – случайности.

Другая общая проблема для агент-каузального либертарианства и либертарианства Франклина – сама концепция агента. Агент-каузальное либертарианство предполагает, что агент – это субстанция. Это, пожалуй, наименее удачное решение, которое достаточно критиковалось в истории философии. Идея же Франклина об отождествлении аген-

та с метажеланием кажется мне контринтуитивной. Даже если допустить, как предлагает Франклин, что метажелание, отождествляемое с агентом, имеет множественную реализацию, сомнительно, чтобы агента представляла лишь одна какая-то установка. Это кажется очень узкой базой для функциональной редукции агента. Такая трактовка не дает основания для достаточно четкого различения и индивидуализации агентов и их состояний, поэтому она вряд ли подходит для данных обстоятельств. Для преодоления проблемы случайности в индетерминистических сценариях необходимо представить более состоятельный анализ концепции агента и его самоопределения в ситуации выбора. В качестве фундамента для такой концепции я предлагаю использовать разновидность нарративного подхода к идентификации личности.

Сторонники нарративной концепции считают, что личность может быть редуцируема к цельной, связной, последовательной истории, биографическому рассказу. «Пожалуй, ядром нарративного подхода является... утверждение, что наши личности по сути своей являются нарративными сущностями» [Schechtman 2011, 395]. Это утверждение разделяют философы А. Макинтайр, Ч. Тейлор, Д. Деннет, М. Шехтман. Связь свойств, событий и действий основывается на возможности непротиворечивого включения их в биографическую историю личности. То есть психологические характеристики, ментальные события и физические действия

принадлежат личности и являются ее частью тогда, когда они могут быть включены в ее биографическую структуру. Биографический рассказ, нарратив, охватывает целую жизнь личности и распределен во времени. Его части расположены в разных временных эпизодах, как части материальных объектов расположены в различных пространственных местах. Значение и осмысленность поступки приобретают только в контексте нарратива. Вне нарратива не существует действия и рациональной личности.

Рациональная личность, в отличие от трехмерных объектов, не присутствует целиком в одном моменте времени, а имеет четыре измерения: три пространственных и одно временное. Агент, взятый в какой-то отдельный момент, – только часть этой протяженной во времени сущности. Такую теорию существования в философии называют *пердурантизмом*, в противоположность эндурантизму – теории, согласно которой объект целиком представлен в один момент времени. Эндурантисты описывают, какие свойства объект имеет в прошлом, настоящем, будущем. Пердурантисты – какие свойства объект имеет безотносительно ко времени. Пердурантизм и нарративная концепция личности, по моему мнению, могут предложить надежный критерий тождества личности во времени и предоставить основание моральной ответственности. Они, в частности, позволяют решить вопрос с ответственностью в случае с индетерминистическим выбором, как он представлен в Аргументе исчезающего агента.

Вот моя интерпретация случая с Ральфом.

Ральф выбирает между отъездом в Нью-Йорк и альтернативой остаться в Мейберри. У него есть мотивы в пользу обоих решений, и выбор его индетерминирован. Альтернативы действий исключают друг друга, а мотивы обладают одинаковой силой. Антецедентные события, являющиеся каузальными факторами, не могут окончательно определить исход, и если ситуация будет воспроизведена множество раз, то выбор Ральфа от случая к случаю будет различным. То есть каузальные факторы, желания и убеждения Ральфа не устанавливают, какой выбор он сделает, а только определяют вероятность принятия каждого из возможных решений равной 0,5. Но, согласно предложенной интерпретации, сам агент, Ральф как пердурантная сущность, *устанавливает*, какой выбор будет сделан, так как выбор конкретной альтернативы является частью (гетерофеноменологического) нарратива обо всей жизни Ральфа, включающего в себя и эпизод с этим самым выбором. Если бы была другая рациональная личность, был бы другой выбор. Поэтому Ральф обладает достаточной степенью контроля, необходимой для моральной ответственности в смысле базовой заслуги.

Нарратив играет роль определяющего фактора по отношению к действию. Это – самоопределение, так как речь идет об определении целым своих частей. При этом в данном случае самоопределение не имеет каузального характера (как считает Франклин), ведь причины, согласно традици-

онной точке зрения, должны предшествовать действию. Отношение между нарративом и выбором – это, скорее, отношение *реализации*. Биографический нарратив реализуется в выборе агента, то есть состоит из множества принятых решений и действий. В свою очередь, нарративный агент *супервентен* на действиях и принятых решениях. То есть невозможно различие личностей в отсутствие различия их решений и действий. Ральф, оставшийся в Мейберри, и Ральф, уехавший в Нью-Йорк, – это две разные личности.

Нарративы, представляющие каждую из них, различаются и, в свою очередь, определяют выбор. Соответственно, каждая из таких личностей несет ответственность за соответствующий выбор, входящий в конкретный нарратив. Таким образом, нарративы представляют условия для контрастного объяснения выбора. А это, как я уже писал, является критерием способности самоопределения и установления выбора. Следовательно, данное дополнение (которое я бы назвал Аргументом четырехмерного агента) снимает потенциальное возражение о случайности некоторых индетерминированных действий, строящееся на основе Аргумента исчезающего агента, и возражение о невозможности моральной ответственности в условиях индетерминизма. Насколько сильна такая позиция?

Защитники аргумента Перебума могли бы попытаться доказать, что нарративный подход не является лучшей концепцией личности или что личность не является пердурант-

ной сущностью. Это, конечно, важные вопросы. Существуют альтернативные подходы к концепции и идентификации личности: психологический, биологический, субстанциальный. Также часть философов считает, что личность не представляет собой пердурантную сущность и что таких сущностей вообще нет. Но это – тема для отдельной дискуссии. Здесь я не берусь защищать эти тезисы. Укажу лишь на то, что среди современных философов есть сторонники предложенных мной взглядов. Достаточно допустить истинность нарративной концепции личности, чтобы признать Аргумент исчезающего агента ложным. Тем самым мяч переходит к Перебуму – это ему теперь необходимо доказать, что нарративная теория ложна, чтобы показать, что его аргумент работает.

Критики моей позиции также могли бы указать на то, что понятие «устанавливать» не подходит для характеристики связи между пердурантной сущностью и спорным выбором: во-первых, потому что пердурантная сущность не антецедентна выбору и, во-вторых, потому что каузально с ним не связана. В защиту Аргумента четырехмерного агента следует указать на то, что такая характеристика не противоречит интерпретации Перебума. В статье «Ответы Тадросу, Смилански, Маккенне и Мили» Перебум разъясняет, что синонимом глагола «устанавливать» в своем аргументе он считает слово «определять». «Устанавливать, какое действие произвести, означает определять, *не обязательно каузально*, ка-

кую из альтернатив действий выбрать», – пишет Перебум (курсив мой. – Д. В.) [Pereboom 2017, 630]. То есть сам автор не ограничивает возможность установления выбора только каузальными факторами. А понятие «определять» подходит для характеристики связи между пердурантной сущностью и ее темпоральной частью. Также в указанной статье автор соглашается, что требование, заключающееся в том, чтобы устанавливающие факторы предшествовали выбору, излишне. «Мили подсказывает, что ошибочно утверждать, что событийно-каузальное либертарианство допускает, чтобы только *предшествующие* выбору события устанавливали, какое решение будет сделано... И я счастлив принять рекомендацию Мили и снять требование о [возможности контроля только над] будущим» [Pereboom 2017, 632–633]. Таким образом, кажется, что моя интерпретация понятия «устанавливать» соответствует представлениям Перебума. А это значит, что в моем аргументе понятия не подменяются.

Критик предлагаемого аргумента мог бы возразить, что описанные обстоятельства все равно не обеспечивают возможность моральной ответственности в смысле базовой заслуги. Это связано с тем, что агент как нарративная пердурантная сущность устанавливает выбор в отношении конкретного действия, но не устанавливает нарратив в целом. То есть сам биографический нарратив является предустановленным для данного агента, и потому агент не может нести за него ответственность. Эта линия критики может быть отне-

сена к традиции критики с помощью Аргумента манипуляций. Как было показано выше, эта линия критики тоже несостоятельна. Таким образом, я считаю доказанным то, что свобода воли и моральная ответственность совместимы как с детерминизмом, так и с индетерминизмом, и, соответственно, суперкомпатибилизм верен. Это – важное достижение, так как суперкомпатибилизм означает, что вопросы о свободе воли и ответственности не являются эмпирическими и целиком находятся в проблемном поле философии. Возможность свободы воли и ответственности не зависит от открытий нейронауки или физики.

V. Итоги второй главы

В начале главы я показал, что рассуждения о наличии или отсутствии у агентов свободы иногда оказываются в тупике: интуиции расходятся и не позволяют сделать на этот счет однозначного вывода. Однако есть способ прояснить эти интуиции. Большинство современных аналитических философов делают это с помощью дополнительного ориентира – ответственности. Идентифицировать случаи наличия или отсутствия ответственности кажется проще. В свою очередь, моральная ответственность служит индикатором наличия необходимого контроля над действием у агента. А этот контроль свидетельствует о наличии метафизической свободы. Таким образом, из рассуждений, приведенных в главе, следует, что свобода и ответственность неразрывно связаны.

Ответственность является важнейшей частью практики взаимоотношений между людьми. Но у рациональных оснований свободы и ответственности есть теоретическое препятствие – детерминизм. Каузальный детерминизм плохо согласуется с двумя принципами, на которых построены свобода и ответственность: Принципом альтернативных возможностей и Принципом автономии. Это следует из Аргумента последствий и Аргумента манипуляций. Впрочем, опасность для свободы и ответственности может представлять также индетерминизм. Это следует из Аргумента случайности. Та-

ким образом, свобода и ответственность могут оказаться в опасности, каким бы образом ни был устроен наш мир.

Взгляды аналитических философов конца XX – начала XXI в. на эти проблемы я представил на примере двух авторитетных инкомпатибилистских позиций: либертарианской теории Р. Кейна и жесткого инкомпатибилизма Д. Перебума. Сходство этих теорий в том, что они отрицают совместимость свободы с детерминизмом. Теории отличаются в том, как интерпретируют соотношение свободы и индетерминизма. Кейн считает индетерминизм необходимым условием свободы и ответственности и в качестве парадигмы ответственного действия приводит индетерминированные волеустанавливающие действия. Перебум отрицает совместимость свободы с индетерминизмом. И в качестве основания для этого приводит Аргумент исчезающего агента. Перебум отрицает также необходимость альтернативных возможностей для свободы на основании примеров «в стиле Г. Франкфурта». Кейн, со своей стороны, альтернативные возможности считает необходимыми для свободы и ответственности. Однако обе эти позиции, с моей точки зрения, ошибочны.

Я показал, что Аргумент последствий, наиболее известный аргумент инкомпатибилистов против совместимости свободы и детерминизма, не убедителен. В этом аргументе понятие «иметь возможность» используется слишком узко, и только на этом искусственном значении понятия строятся выводы. Я также попытался показать, что Аргумент мани-

пуляций в формулировке Перебума, целью которого также является демонстрация несовместимости детерминизма со свободой и ответственностью, на самом деле вызывает противоположные интуиции. Это становится лучше видно с помощью Сценария «Радиоуправляемый герой». Аргумент исчезающего агента, призванный, в свою очередь, показать несовместимость свободы воли с индетерминизмом, я попытался отвести через контраргумент четырехмерного агента. Следствием критики инкомпатибилистских аргументов стало обоснование суперкомпатибилизма, согласно которому свобода и ответственность совместимы как с детерминизмом, так и с индетерминизмом.

Тем не менее осталась как минимум еще одна важная нерешенная проблема: проблема связи личности с действием. Без ее решения рациональных оснований для признания возможности ответственности и свободы все равно недостаточно. Чтобы решить эту проблему, потребуется обратиться к теориям личности и найти наиболее правдоподобную среди них. К анализу современных теорий личности я перейду в третьей, заключительной главе этой книги.

Глава 3. Проблема тождества личности и моральная ответственность

I. Событийный каузализм как угроза тождеству личности

1. Связь вопроса о тождестве личности с проблемами моральной ответственности и свободы

Во второй главе я решал проблему совместимости свободы с детерминизмом. В третьей главе я собираюсь изучить вопрос о совместимости свободы воли, моральной ответственности и событийного каузализма. В первой части главы я объясню, что следует называть событийным каузализмом и почему эта концепция важна для современной метафизики. Потом я покажу, почему событийный каузализм представляет препятствие для свободы воли и моральной ответственности и ведет к проблеме асинхронного тождества личности. Таким образом, в центре внимания в данной главе

будут проблемы существования личности и ее асинхронного тождества. В поисках решения этих проблем я обращусь к основным классическим подходам современных аналитических философов: Д. Перри и С. Шумейкера, Э. Олсона и Р. Суинбёрна.

Сначала я расскажу про психологический подход и его истоки в теории личности Дж. Локка. Потом я представлю развитие этой теории в концепциях Перри и Шумейкера. Я объясню, каким образом современный психологический подход позволяет связать стадии личности на разных временных этапах в одну последовательность. Далее я представлю аргумент Б. Вильямса и мысленный эксперимент «Теле-транспортёр» Д. Парфита. Я считаю, что именно эти рассуждения показывают основную уязвимость психологического подхода – проблему редупликации личности.

В следующем параграфе будет описан подход Э. Олсона, предлагающего биологические критерии тождества личности. Я покажу, что постановка проблемы в биологическом подходе ведет к отдельным трудностям. Олсон начинает решать проблему реидентификации личности, но по факту идентифицирует личность и объект. Такое решение не кажется мне адекватным исходной проблеме. После критики биологического подхода я обращусь к субстанциональной теории Р. Суинбёрна. Она призвана решить трудности, возникающие в рамках психологического и биологического подходов. Но и здесь, по моим представлениям, не обходится

без проблем. Основной проблемой субстанциональной теории является ее несовместимость с Принципом верификации. Общий вывод о недостатках классических подходов к проблеме тождества личности будет резюмирован Аргументом пяти трансплантаций. На этом закончится критическая часть главы. Вторая ее часть будет позитивной.

Наиболее перспективным решением проблемы тождества личности, по моему мнению, является нарративная теория личности. Это относительно новый, неклассический подход. В нем успешно решен вопрос о совместимости событийного каузализма и свободы, а также обнаруживаются надежные основания для приписывания моральной ответственности личности за ее поступки. В ходе исследования я выделю общие элементы в теориях сторонников нарративного подхода, подчеркну его преимущества перед альтернативными подходами и попытаюсь ответить на ряд имеющихся возражений.

Отдельное внимание я уделю концепции Д. Деннета и его теории личности как центра нарративной гравитации. В текстах Деннета нарративный подход представлен в основном в виде мысленных экспериментов и художественных иллюстраций. Это вынуждает провести их подробный анализ и попытаться расшифровать ключевые идеи предлагаемого им решения. Так, я проанализирую его мысленный эксперимент «Где Я?». Этот анализ позволит лучше понять критерии тождества личности, предлагаемые американским философом.

Далее я покажу эффективность работы этих критериев не только для обычных, но и для патологических случаев.

В заключительной части третьей главы я представлю дополнительные доводы в пользу нарративного подхода. Я обращаюсь к идеям нейрофизиологов и психологов Д. Вегнера, М. Газзаниги и М. Тейлор. Проведенный ими анализ различных патологий и особенностей развития психики (диссоциативного расстройства личности, работы разделенных полушарий мозга, а также феномена «вымышленных агентов») позволяет укрепить убеждение в нарративной природе личности. Таким образом, результатом третьей главы будет обоснование нарративного подхода и концепции личности как центра нарративной гравитации.

2. Событийный каузализм

Каузализм – это убеждение в том, что у каждого события есть причина. Каузализм в некотором смысле является синонимом детерминизма, а также обычно выражает натуралистическое представление о природе действия. Это понятие активно использует А. Мили в статье «Каузальность, действие и свобода воли» [Mele 2009]. Стандартным каузализмом Мили называет теорию действия Дэвидсона, о которой мы говорили в первой главе. Событийным каузализмом я предлагаю называть теорию действия, в которой причинами действий (или любых других событий) могут быть только

события.

Этот вид каузализма совместим как с позицией Дэвидсона, так и с позициями его оппонентов. Но он несовместим с позицией агентного каузализма, то есть с убеждением в том, что не только события, но и сами агенты как субстанции могут выступать причинами действий. Современными агентными каузалистами являются некоторые либертарианцы, такие как К. Лерер, Т. О'Коннор, Р. Кларк. Тем не менее эта позиция является одной из наиболее экстравагантных и не находит большой поддержки у современных философов. Общепринятой теорией является, наоборот, событийный каузализм. Это подтверждает Д. Эринг в статье «Реляты каузального отношения»: «Согласно наиболее популярному взгляду, релятами каузального отношения являются неповторяющиеся события, включая изменения, которые мы обычно считаем событиями, но также включая и неизменные состояния» [Ehring 2009, 388].

Тем не менее событийный каузализм влечет за собой проблему. Этой проблемой является Исчезновение личности. У нее – два аспекта. Во-первых, когда только события (пусть даже ментальные) являются причинами других событий, остается непроясненным, какое отношение действие имеет к агенту, какую роль он играет по отношению к действиям. Беспокойство об этом аспекте выражает, в частности, философ сознания Т. Нагель: «Важный источник проблемы – это представление о личностях и их действиях как

части природного порядка... Эта концепция... ведет к ощущению, что мы – вообще не агенты... Кажется, что нет места агентам в [таком] мире» [Nagel 1986, 110–111]. Отсутствие каузально эффективного агента может казаться основанием отрицать существование агента вообще. Но если агента не существует, кто может обладать свободой и кто должен нести моральную ответственность?

Вторым аспектом этой проблемы является возможность длительного существования агента. Согласно общим представлениям, чтобы нести ответственность за действие, агент должен находиться в каком-то очень близком отношении к этому действию. Общепринятым считается положение, что агент «может нести моральную ответственность только за собственные действия». Это утверждение фиксирует Д. Шумейкер в статье «Моральная ответственность и личность» [Shoemaker 2011a, 488]. Но что подразумевает это убеждение? Для многих теоретиков свободы воли и моральной ответственности оно означает, что «моральная ответственность предполагает тождество личности [во времени]» [Shoemaker 2011a, 488]. Иными словами, чтобы агент, совершивший действие в момент t_1 , мог нести за него ответственность в момент t_x , он должен быть тем же самым агентом в этот момент. Но в условиях событийного каузализма возникновение и длительное существование тождественного агента проблематично. Это означает, что событийный каузализм, наряду с физикализмом и детерминизмом, представ-

ляет новую опасность для свободы и ответственности. «Чтобы прояснить ключевые компоненты моральной ответственности, – пишет Шумейкер, – нам нужно исследовать природу тождества личности» [Shoemaker 2011a, 488]³⁶.

Вопрос о моральной ответственности, таким образом, оказывается тесно связан с вопросом о тождестве и единстве личности. Ответственность и свобода не могут существовать без тождества личности, а любая теория тождества личности должна решать вопрос об ответственности. С этой позицией согласен Д. Парфит. Он утверждает: «Некоторые важные вопросы [о том, что необходимо для выживания, воспоминаний и ответственности] действительно предполагают вопрос о тождестве личности. Но их можно освободить от этой предпосылки. И когда это будет сделано, вопрос о тождестве перестанет быть значимым» [Parfit 1971, 4]. Вполне вероятно, что существуют какие-то другие смыслы тождества личности, но споры о них могут оказаться совсем пустыми. Именно тот, кто несет ответственность за мои поступки в прошлом, является тождественным мне в практически значимом смысле.

³⁶ С этим утверждением согласна, в частности, М. Секацкая: «Уйти от детерминизма и прийти к свободе невозможно без носителя свободы воли, без субъекта, свободным волеизъявлением предопределяющего собственные действия» [Секацкая 2012, 90]. А также С. М. Левин [Левин 2013].

3. Проблема единства и тождества личности

Проблема единства и тождества личности – классическая философская загадка. Она может быть обнаружена при сопоставлении обыденных представлений и философского анализа личности. Согласно обыденным представлениям, личность – это единая конкретная сущность. У нее есть пространственное расположение – место, в которое, как в «командный пункт», стекается вся информация из внешнего мира, поступающая через органы чувств, где она объединяется и где выносятся решения действовать. Все действия личности являются результатом решений, поступающих из этого «командного и волевого центра». Это – первая интуиция. Согласно другой, предполагается, что в обычных условиях с течением времени личность остается той же самой. Личность существует на протяжении жизни человека, и в разные моменты времени это одна и та же личность. Личность может сохраниться или исчезнуть, но исчезла она или сохранилась – это объективный факт. И наконец, согласно третьей интуиции, кажется, что личность – это что-то уникальное. Может существовать только одна *такая-то* личность. Личность не может размножиться. То есть личность – это нечто простое, целое, неделимое, уникальное и дискретное. Кажется непредставимым, чтобы некоторый чело-

век был частично «мною» и «не мною». Допустимо, чтобы мы не знали, что личность X в прошлом – это личность Y в настоящем. Но недопустимо, чтобы это было принципиально неопределимо. Эти идеи содержатся в обыденном представлении о личности.

С другой стороны, поиски простой и неделимой сущности кажутся обреченными. При философском и научном исследовании устройства человека выясняется, что оно отчасти напоминает часовой механизм. В нем есть множество важных частей, которые выполняют каждая свою функцию, и нет главной точки, единого физического центра личности. Нет «точки сбора» ментальных содержаний. О том, что *собственное Я* нельзя обнаружить, наблюдая человека через увеличительные приборы, размышлял Лейбниц еще в начале XVIII в. Он предлагал вообразить исследователя входящим в «машину», производящую мысль и восприятия. Во время осмотра этой «машины», говорил он, нельзя обнаружить ничего, кроме частей, толкающих друг друга. И трудно представить, как взаимодействие этих частей приводит к мыслям, восприятиям и решениям [Лейбниц 1982, 415].

Сегодня убеждение в отсутствии единой точки контроля и волевого управления получило дополнительные нейрофизиологические подтверждения. По мнению современных нейрофизиологов, Я – «это набор параллельно работающих и распределенных по разным частям мозга процессов без центрального пункта управления» [Gazzaniga 2011, 107]. В

мозге есть тысячи мест, где принимаются решения и одновременно совершается выбор. Но в мозге, в этой биологической машине, нет духа, нет субстанционального Я. Это – первая значительная проблема, проблема синхронного единства.

Трудности возникают и при попытке обнаружить длящуюся во времени часть личности. Клетки организма регулярно обновляются: появляются новые, старые отмирают. За несколько лет обновляется весь организм. Что тогда составляет эту неизменную часть? Найти длящуюся часть личности нельзя как объективными методами, так и с помощью интроспекции. Стоит вспомнить рассуждения Юма о поиске Я во внутреннем опыте: «Что касается меня, то, когда я самым интимным образом вникаю в нечто, именуемое мной своим я, я всегда наталкиваюсь на то или иное единичное восприятие тепла или холода, света или тени, любви или ненависти, страдания или наслаждения. Я никак не могу уловить свое я как нечто существующее помимо восприятий» [Юм 1996а, 298]. Восприятия сменяют друг друга, мысли и воспоминания следуют поочередно, и в них нельзя наблюдать, с точки зрения Юма, что-то, что сохраняло бы тождественность в ходе этих изменений. Однако такая картина не согласуется с представлением о моральной ответственности: без единой длящейся во времени личности ответственность невозможна. Юм сам признавался позже в неудовлетворительности своих выводов и отказывался от обозначенной позиции

[Юм 1996а, 325–326]. Таким образом, Проблема исчезающей личности и вопрос о том, что обеспечивает тождество личности во времени, сохраняют свою актуальность.

II. Современные классические теории тождества личности

Каковы необходимые и достаточные условия того, чтобы один человек, одна личность в момент t_1 был тем же человеком, той же личностью, что и в момент t_2 ?³⁷ В этом вопросе свойство быть «тем же» может обозначать две разные характеристики: нумерическую и качественную тождественность. Когда говорят о нумерическом (метафизическом) тождестве вещей, имеют в виду, что эти две вещи являются фактически одной, одной и той же вещью. Когда говорят о качественной тождественности, говорят, что вещи полностью схожи, абсолютно одинаковы. Но качественно тождественных вещей может быть нумерически две или больше. Две элементарные частицы могут быть абсолютно качественно тождественны, что означает, что они будут полностью идентичны по своим свойствам и проявлениям, но они будут двумя отдельными частицами. При этом один и тот же человек или одна и та же вещь в разное время может быть качественно различной, но нумерически тождественной. Так, работающие и сломанные часы различаются между собой качественно, но могут являться одними и теми же часами в разные моменты вре-

³⁷ Аналогичным образом проблему асинхронного тождества личности формулирует Д. Парфит [Parfit 1986, 202].

мени. Какой из вариантов тождественности необходим для атрибуции моральной ответственности, если необходимо вообще, нам еще предстоит выяснить.

Современные аналитические философы в большинстве своем разрабатывают три классических подхода к вопросу о единстве и тождестве личности. Это субстанциональный, психологический и биологический подходы. Психологический и биологический являются *реляционными подходами*. Они предполагают, что между личностью в момент t_2 и личностью в момент t_1 существует какое-то отношение. Реляционные подходы отличаются от более экзотического для философского сообщества и наиболее уязвимо для критики *субстанционального подхода*. Но именно субстанциональный подход к тождественности личности распространен больше всего за пределами профессионального дискурса. Субстанциональный подход предполагает, что тождественность между двумя фазами личности обуславливается тем, что обе фазы реализуют одну и ту же субстанцию. Реляционные подходы являются *редукционистскими*, так как вопрос о тождестве личности они определяют на основе фактов о мозге, физическом теле, связи физических и ментальных событий. Нередукционистским является субстанциональный подход, так как вопрос о тождестве личности (трактуемой обычно как картезианское Я или душа) он не ставит в зависимость от «прочих фактов», связанных с физическими или ментальными событиями. Все описанные выше подходы

представляют способы отождествления личности в различные моменты времени и направления решения обозначенного выше вопроса реидентификации.

1. Психологический подход

В аналитической философии конца XX – начала XXI в. доминирует психологическая теория³⁸. Этот историко-философский факт подтверждает, в частности, М. Шехтман. Она пишет: «На протяжении последней половины столетия имелся значительный интерес к разработкам и защите психологических подходов к тождеству личности... Существуют десятки версий теории непрерывности психологических черт и множество талантливых сторонников этого общего подхода» [Schechtman 2014, 20–21]. Согласно психологическому подходу, тождество личности определяется психологическими характеристиками. Личность продолжает существовать тогда, когда происходит наследование психологических характеристик: воспоминаний, убеждений, желаний, характера; и исчезает, когда наследников этих качеств не остается. Одним из первых сторонников этого подхода был Дж. Локк.

³⁸ В опросе, проведенном в 2009 г. среди 900 англоговорящих профессоров философии, 34 % утверждали, что разделяют тот или иной вариант психологического подхода, 17 % – физиологический подход, 12 % – субстанциональный. Остальные 37 % не имели определенного мнения [Kind 2015, 102].

А. Дж. Локк: память как критерий тождества личности

С именем Локка связывают в целом постановку проблемы тождества личности в современной философии. В «Опыте о человеческом разумении» он формулирует проблему и предлагает собственное решение, которое можно считать исходной психологической теорией тождества личности. В этой теории ключевым критерием тождества является преимущественно память³⁹: «...так что все то, что имеет сознание настоящих и прошедших действий, есть одна и та же личность, к которой относятся и те и другие действия» [Локк 1985, 393]. Локк считал, что существование личности простирается в прошлое настолько, насколько прошлое может быть доступно в воспоминаниях, и в будущее настолько, насколько будет распространяться индивидуальное сознание. Этот подход обладал множеством достоинств.

Прежде всего, кажется, что именно на памяти основано представление человека о длительности собственного существования. Мы помним события прошлого и из этого, ви-

³⁹ Концепция Локка на самом деле может быть интерпретирована различными способами. Большинство современных аналитических философов считает, что Локк выделял именно память как основу тождества личности. Однако в текстах Локка в качестве основы тождества часто фигурируют также сознание и психологические характеристики. В данном случае, для целей нашей работы, я считаю возможным некоторое упрощение этой теории.

дим, заключаем, что присутствовали в прошлом. Помимо этого, подход Локка построен на интуитивно достоверном убеждении, что сущность человека заключается в его личностных психологических качествах. «Для того чтобы выяснить, в чем заключается тождество личности, – пишет Локк, – мы должны рассмотреть, что такое “личность”; личность, я думаю, – это мыслящее, разумное существо, имеющее разум и рефлексию, которое способно считать себя собой, одним и тем же мыслящим существом в разное время и в разных местах. И это оно делает лишь посредством сознания, которое неотделимо от мышления» [Locke 2008, 39]. Приоритетное значение психологических свойств, в частности памяти, в вопросе об идентификации иллюстрируется гипотетической ситуацией о принце и сапожнике.

Локк предлагает вообразить ситуацию, в которой принц и сапожник меняются воспоминаниями: все воспоминания принца передаются сапожнику, а мысли и воспоминания сапожника – принцу. Кто из них кто? – спрашивает философ. Разумное большинство сочтет, что принц оказался в теле сапожника, а сапожник – в теле принца. Из этого утверждения Локк делает вывод, что идея тождества памяти лежит в основе обыденных интуиций о тождестве личности. Именно память, а не тождество, например, тела, определяет, кто есть кто.

Подтверждением приоритетности психологических критериев также является особая озабоченность многих людей

существованием после физической смерти. Этому не мешает знание о том, что биологическая жизнь конечна и что тело рано или поздно перестанет существовать. Пусть человеческое тело погибает, уничтожается, считают они, но личность все равно может быть сохранена. Интуитивно кажется, что для продолжения существования важно сохранение сознания, мышления, воспоминаний и других психологических характеристик. Наличие идентичного тела при этом не играет существенной роли.

Идеи Локка согласуются с обыденными интуициями и в других случаях, например в отношении поступков, совершенных в состоянии невменяемости. Так, его психологический подход к идентификации объясняет, почему личность не должна нести полной ответственности в таких случаях. Интуитивно кажется, что сумасшедший человек – не тот же самый человек, что человек здравомыслящий. И критерий доступности в памяти это вполне оправдывает. Поступки, совершенные в состоянии невменяемости, часто недоступны памяти человека. А значит, согласно психологическому критерию, их нельзя приписывать той же личности. Здесь вполне уместно говорить: «он был сам не свой». Невменяемый человек, в соответствии с психологическим критерием, – не та же самая личность, что человек здравомыслящий, даже если они представляют одно и то же биологическое существо. А значит, на этом основании здравомыслящего человека можно освобождать от ответственности за поступки

человека сумасшедшего.

Описанный выше подход дает средства связать между собой различные асинхронные «фазы» существования личности и намечает путь синхронного объединения личности. Память является одним из компонентов сознания. А сознание может являться основанием для синхронного объединения личности. Таким образом, подход Локка заслуживает внимания и перспективен. Но решение это в своем исходном варианте все-таки неверное. Уже ближайшие оппоненты обратили внимание на явные недостатки памяти как критерия тождества личности и атрибуции поступков. Дж. Батлер и Т. Рид, сторонники альтернативного, субстанционального критерия, выступили с острой критикой исходного варианта психологического подхода.

В. Возражения к теории памяти Дж. Локка

Батлер и Рид разделяли убеждение Локка в том, что для ответственности нужна тождественность личности. В частности, Рид писал, что «тождественность [личности]... является фундаментом всех прав и обязанностей...» [Reid 1969, 112]. Но оба философа были полностью не согласны с его основанием тождества. Батлер указывал, что нельзя принимать сознание за критерий тождества, так как содержание сознания постоянно изменяется. По мнению Батлера, из идей Локка следует, что никто не может отвечать за «свои» по-

ступки в прошлом, так как никто не является тем же человеком сегодня (в плане сознания), что был вчера. А это следствие противоречит обычной практике, что является основанием для сомнений. Помимо этого, Батлер находил *логический круг* в локковских построениях. По его мнению, доступность сознанию каких-то действий в воспоминаниях уже предполагает установление тождественности личности, совершавшей действия, и потому сознание и память не могут быть основанием тождественности. То есть никто не может действительно *помнить* поступки, предварительно не идентифицировав как-то совершающего их с самим собой.

Предположим, мы хотим узнать, является ли Блотт, который существует сейчас, Клоттом, о котором мы знаем, что он существовал раньше. Из критерия Локка следует, что Блотт – это Клотт, если Блотт может в памяти иметь доступ к состояниям Клотта, которые имели место в прошлом. Но «воспоминания» Блотта о состояниях Клотта будут только в том случае воспоминаниями, если Блотт – в действительности Клотт. Поэтому, чтобы использовать критерий памяти, нам нужно сначала идентифицировать Блотта с Клоттом. Конечно, Блотт может воображать или ошибочно думать, что он «вспоминает» состояния Клотта. Но такие воображаемые или ошибочные «воспоминания» не будут действительно воспоминаниями. Это будут иллюзии или фантазии. Следовательно, не память делает поступки собственными, а что-то еще.

Батлер считал позицию Локка недоразумением. А Рид дополнял критику яркими контрпримерами. Он обращал внимание на то, что у этой теории – странные, подозрительные следствия: так, согласно теории Локка, один и тот же человек может быть и не быть субъектом одного и того же действия. Рид демонстрирует это с помощью гипотетической ситуации с отважным офицером. Предположим, храбрый офицер в детстве воровал яблоки из чужого сада. Потом он похитил вражеский флаг во время своей первой военной кампании, а еще позже получил звание генерала. Предположим также, что молодой офицер, когда похищал вражеский штандарт, помнил, как воровал яблоки в чужом саду, а когда получал звание генерала, еще помнил, как похищал флаг, но не помнил и не мог помнить воровства яблок в детстве. Если такая ситуация возможна, то, согласно теории Локка, человек, похитивший штандарт, – та же личность, что ребенок, воровавший яблоки, а генерал – тот же человек, что похищал вражеский штандарт. То есть старый генерал – это храбрый офицер, а офицер – озорной ребенок. Но по критерию памяти старый генерал – не тот же человек, что озорной ребенок! Выходит, генерал одновременно и тот человек, и не тот человек, который воровал яблоки [Reid 2008, 115].

Гипотетическая ситуация вскрывает важное различие между отношением тождества и отношением доступности в памяти. Если x тождественен y , а y тождественен z , то x должен быть тождественен z . То есть тождественность – это

транзитивное отношение. Но если x помнит то, что происходило с y , а y помнит то, что происходило с z , это еще совсем не значит, что x помнит, что происходило с z . То есть доступность некоторого воспоминания в сознании – не транзитивное отношение. И это – явная проблема для психологического критерия, и еще не последняя проблема.

Интуитивно кажется, что тождество личности сохраняется в случае полной потери памяти. Даже в таком случае человек продолжает существование, оставаясь той же личностью. Однако, по локковскому критерию, человек, потерявший память, – не тот же человек, у которого были воспоминания, поскольку сознание первого будет существенно отличаться от сознания второго. Как можно парировать все эти возражения?

С. Современные вариации психологического подхода

Среди современных аналитических философов сторонниками психологического критерия тождественности личности являются Э. Кинтон, П. Грайс, Дж. Перри и отчасти С. Шумейкер и Д. Парфит. Их задачей является усовершенствование идей Локка, разрешение трудностей, обнаруженных критиками. Сторонники психологической теории, на мой взгляд, должны (1) преодолеть обвинения в цикличности критерия тождества, (2) решить проблему транзитив-

ности и (3) найти способ сохранять тождество личности в случаях потери памяти и сознания. Попытки преодолеть эти препятствия ведут к формулировке более сложных психологических критериев тождества.

Проблему (1) цикличности памяти как критерия тождества личности попытались решить С. Шумейкер [Shoemaker 1970] и Д. Парфит [Parfit 1986]. Их предложение состояло в том, чтобы ввести новое, техническое понятие квазипамяти, или *q*-памяти, которое бы не предполагало тождественности помнящего и пережившего события в прошлом. Так, я могу иметь *q*-воспоминания о переживаниях в прошлом в том случае, если эти переживания были у *кого-то* в прошлом и если мои воспоминания каузально связаны правильным образом с этими переживаниями. Тогда обычная память будет просто подклассом *q*-памяти. А *q*-память, в свою очередь, можно использовать для теории тождества личности, избегая цикличности.

Не менее элегантное решение было найдено и для проблемы (2) транзитивности (в асинхронной ситуации). Оно состояло в объединении фаз личности с помощью пересекающихся воспоминаний или *преемственности (continuity)*. В отличие от памяти преемственность является транзитивной связью. Действительно, воспоминания о событиях далекого прошлого со временем угасают. Зато сравнительно недавние события помнятся отчетливо. Генерал помнит события, которые произошли с отважным офицером, офицер помнит

события, которые произошли с озорным ребенком. И генерал может не помнить события, которые произошли с ребенком. То есть между генералом и ребенком *нет прямой связи через память (direct memory connections)*. Но генерал связан с ребенком через множество промежуточных фаз, между ними есть преемственность. Воспоминания генерала сейчас и в предыдущий день большей частью пересекаются, аналогичным образом они пересекались вчера и позавчера, и так – до самого детства. Допустим, условно, что в детстве ребенку были доступны воспоминания 1–5, во время службы офицером – с 3 до 8, во время отставки в звании генерала – с 6 до 10. Тогда между всеми фазами ребенка-генерала всегда было значительное количество пересекающихся воспоминаний (в нашем случае – 3). А это значит, что между генералом и ребенком существует косвенная связь через память, или *преемственность памяти (continuity of memory)*. Эту преемственность и можно взять за основу тождественности. В таком случае проблема с транзитивностью и контрпримерами вроде случая с «храбрым офицером» будет решена.

Третью проблему можно попытаться обойти, введя дополнительные, более широкие критерии тождества. Помимо памяти, можно опираться на другие психологические свойства, такие как намерения, реализуемые в действиях, цели, убеждения, желания и характер. Тогда даже в условиях потери памяти, но при сохранении других психологических качеств, личность можно считать сохранившей идентичность.

Суммировав все эти усовершенствования, можно получить новый *психологический критерий тождества личности*. Шумейкер предлагает такую формулировку: « X в момент t_1 будет той же личностью, что Y в момент t_2 , если и только если X имеет уникальную непрерывную психологическую преемственность с Y , выражающуюся во множестве пересекающихся цепочек сильной психологической связности (*connectedness*), заключающейся в большом числе прямых психологических связей, таких как память, намерения, убеждения, цели, желания и общности характера» [Shoemaker 2016]. Этот усовершенствованный критерий, кажется, выдерживает обозначенную выше критику. Но в конечном итоге даже этот новый психологический критерий оказывается неудовлетворительным.

Основной проблемой для современных психологических теорий является проблема редупликации. Мысленный эксперимент, раскрывающий эту проблему, был представлен американским философом Б. Вильямсом. Он является важным вкладом современной философии в исследование вопроса реидентификации. Можно даже утверждать, что после того, как проблема тождества личности была инициирована в философии Локка и развита другими философами Нового времени, «контуры дискуссии вокруг проблемы были зафиксированы и оставались без изменений вплоть до 1956 г. В этом году, однако, Бернард Вильямс опубликовал свою новаторскую статью... в которой он выдвинул свой знаменитый

Аргумент редупликации. Этот аргумент трансформировал последующее обсуждение проблемы и привел философов к формулировке принципиально новых позиций» [Noonan 2003, 125].

Д. Аргумент редупликации и Аргумент телепортации

В статье «Тождество личности и индивидуализация» (1956 г.) Вильямс критикует психологические теории тождества с помощью мысленного эксперимента. Автор предлагает вообразить ситуацию, в которой человек по имени Чарлз, живущий в XX в., утверждает, что он – Гай Фокс, английский дворянин, участник заговора против английского короля в XVII в. Согласно мысленному эксперименту Вильямса, Чарлз знает в точности все детали жизни Фокса, даже неизвестные историкам. При этом все утверждения Чарлза о Фоксе находят удивительные подтверждения, которые могут быть объяснены только тем, что Чарлз действительно помнит жизнь Гая Фокса. Вильямс допускает, что кто-то мог бы интерпретировать этот случай как реинкарнацию, то есть утверждать, что Чарлз является той же личностью, что и Гай Фокс. Но автор обнаруживает, что такое суждение ведет к парадоксу.

«Если логически возможно, чтобы с Чарлзом произошли описанные изменения (то есть чтобы он обрел воспомина-

ния Гая Фокса. – Д. В.), то логически возможно, чтобы другой человек одновременно претерпел такие же изменения; то есть, например, чтобы как Чарлз, так и его брат Роберт оказались в этом состоянии» [Williams 1956–1957, 238]. Тогда следует признать, что оба брата являются той же личностью, что и Гай Фокс. Но это утверждение, по мнению Вильямса, противоречиво. Ибо тогда получается, что Гай Фокс находится в двух местах одновременно и что оба брата как личности тождественны друг другу, а это невозможно.

Из рассуждений Вильямса следует также, что, даже когда Чарлз – единственный, кто помнит жизнь Гая Фокса, оснований считать его Гаем Фоксом тоже недостаточно. Ведь факт тождества Чарлза и Гая Фокса не может зависеть от существования или отсутствия других личностей с какими-то воспоминаниями. Тождественность является внутренним (*intrinsic*), а не контекстуально зависимым отношением. Таким образом, Аргумент редупликации Вильямса подрывает убеждение, что воспоминания и иные содержания сознания могут служить основанием тождественности личности. Аналогичным образом, Аргумент дает основания считать любые психологические критерии недостаточными для тождества личности, ведь эти критерии также могут быть «редуплицированы» Чарлзом и Робертом.

Впрочем, защитник психологической теории мог бы возразить: Вильямс приводит любопытный сюжет, но этот сюжет непрояснен. Философ никак не объясняет, каким обра-

зом могут оба брата, Чарлз и Роберт, приобрести воспоминания Гая Фокса. Парировать это возражение можно с помощью другого мысленного эксперимента, который предлагает Д. Парфит в уже ставшей классической работе «Причины и личности» (1984). Эта фантастическая история о космическом путешествии с помощью телепортации называется «Телетранспортер». Парфит рассказывает ее от первого лица:

Для путешествия с Земли на Марс мне предлагают воспользоваться новым аппаратом для телепортации. Аппарат значительно экономит время и расходы на космические путешествия. Аппарат состоит из двух частей: сканера и репликатора. Сканер на Земле записывает всю информацию о состоянии моего тела и мозга и передает ее на Марс. Затем все мои клетки на Земле мгновенно уничтожаются. Репликатор на Марсе, получив информацию с Земли, из нового материала воссоздает точную копию моего тела. Именно это тело продолжает мою жизненную траекторию на Марсе. Это тело обладает в точности такими же психическими характеристиками, что и уничтоженное тело на Земле. Таким образом, аппарат телепортации обеспечивает психологическую преемственность. Обдумав доводы за и против, я соглашаюсь на телепортацию.

Я захожу в кабину для сканирования, нажимаю большую зеленую кнопку... но ничего не происходит. После короткого ожидания в кабине я обращаюсь к диспетчеру: «У Вас сломался аппарат!» В недоумении диспетчер отвечает: «Одну минуточку, сейчас

проверю». Через некоторое время он возвращается на связь: «Вы знаете, вся информация о составе вашего тела была скопирована и передана на Марс. Там уже собрана ваша точная копия. Правда... процесс уничтожения на Земле не закончился. Вместо полного уничтожения у вас повредилась только сердечно-сосудистая система. Не беспокойтесь, через несколько дней у вас наступит приступ, и ваша копия на Земле будет уничтожена. А на Марсе вы продолжите свою жизнь абсолютно здоровым. Если хотите, я сейчас соединю вас с вашей точной копией и, поговорив с ней, вы убедитесь, что это ваш полный психологический двойник». [Parfit 1986, 199–200].

Сломанный телетранспортер представляет собой вариацию случая редупликации, описанного Вильямсом. Однако в отличие от исходного варианта, мысленный эксперимент Парфита обладает важным достоинством. Он хотя бы в общем виде описывает механизм, с помощью которого личность может быть раздвоена. Представляется, что такой механизм возможен хотя бы в принципе, а значит, логически возможна и описанная ситуация. При этом она, кажется, сводит психологическую теорию тождества к абсурду: у Парфита после телепортации есть двойник. Обе личности обладают психологической преемственностью, выражающейся во множестве пересекающихся сильных психологических связей, таких как память, намерения, убеждения, желания и общность характера, то есть, согласно психологическому крите-

рию, они в равной степени тождественны Парфиту до телепортации. При этом Парфит на Земле и Парфит на Марсе – нумерически две разные личности, не тождественные друг другу. В формальном виде результат выглядит так: a тождественно b , a тождественно c , при этом b не тождественно c . Это новый случай нарушения транзитивности связи (теперь в синхронной ситуации). Такое следствие является серьезной проблемой для сторонников современных психологических теорий тождества личности.

Е. Решение проблемы редупликации в современных психологических теориях

Я могу выделить два способа преодоления проблемы редупликации, доступных для сторонников современных психологических теорий: 1) отрицание возможности реализации случаев, описываемых в мысленных экспериментах с редупликацией; 2) введение дополнительного условия тождества. Первый путь – это решение в духе Витгенштейна и Куайна. Этот путь заключается в отрицании реальной значимости отдельных мысленных экспериментов или метода в целом. Ситуации, которые описывают Вильямс и Парфит, являются вымышленными. Более того, в обоих случаях механизмы их реализации могут оказаться физически неосуществимыми. Вполне возможно, что тогда на них нельзя полагаться вовсе, и нельзя доверять интуициям, возникающим

по отношению к ним.

Этот способ решения проблемы редупликации мне кажется наиболее слабым. Дело в том, что метод мысленных экспериментов активно используется самими сторонниками психологического подхода. Достаточно вспомнить пример с принцем и сапожником Локка. Без опоры на подобные гипотетические ситуации сложно было бы обосновать преимущества психологического подхода по отношению, например, к биологическому или субстанциональному. Кроме того, существуют более реалистичные сценарии редупликации личности, например путем «расщепления» личности в результате пересадки одного полушария мозга в другое тело. В результате такой операции тоже могут появиться две личности с преемственными психологическими характеристиками, при этом они не будут тождественны друг другу. Фактическая возможность подобной операции подтверждается успешным опытом пересадки части мозга у мышей, проведенным учеными в начале 1980-х [Kind 2015, 47]. Таким образом, есть основание считать, что первый путь не является перспективным для преодоления проблемы редупликации в контексте психологических теорий.

Второй способ заключается во введении дополнительного критерия тождества личности – отсутствия двойников. Согласно обновленной теории, личность *Б* является тождественной личности *А* в том случае, если *Б* обладает психологической преемственностью по отношению к *А* и нет другой

личности, которая также обладала бы психологической приемственностью по отношению к А. Результатом такого нововведения, в частности, является отрицание тождества между Парфитом-героем до телепортации и двумя его приемниками. То есть, согласно новому критерию, Парфит-герой перестает существовать в случае редупликации. Это решение снимает проблему отсутствия транзитивности в психологических теориях. Но и у этого решения есть недостатки.

Введение дополнительного критерия, предполагающего отсутствие «расщепления», вероятно, нарушает «Принцип только x и y » [Kind 2015, 61]. Согласно этому принципу, тождество x и y зависит только от фактов об x и y . Как я уже говорил, тождественность является внутренним, а не контекстуально зависимым отношением. «Принцип только x и y » кажется достоверным, а его нарушение ведет к довольно любопытным следствиям. Их замечает Э. Кайнд. Она приводит остроумную вариацию парфитовского сюжета.

Представим, пишет Кайнд, что сломанный аппарат для телепортации создает несколько копий с небольшой задержкой. Парфит-герой заходит в камеру для телепортации на Земле, его тело уничтожается, и точная его копия возникает на Марсе. Но спустя несколько минут появляется еще одна копия. Учитывая дополнительный критерий, пока копия Парфита-героя существует в единичном экземпляре, у него есть тождественная личность. Но как только появляется второй двойник, тождественная Парфиту-герою личность исче-

зает. Но как может двойной успех быть неудачей?

Мне представляется возможным еще больше усугубить абсурдность следствий второго подхода. Для этого стоит представить, что через некоторое время один из клонов погибает. Теперь опять возникает тождественная Парфиту-герою личность. Таким образом, Парфит-герой может существовать, исчезать и появляться вновь сколько угодно раз и с любыми паузами и промежутками во времени. То есть Парфит будет «мерцающей личностью». При этом все это будет абсолютно незаметно как с позиции первого лица, так и с позиции третьего лица. Кажется, что это действительно абсурдное следствие, то есть нельзя сказать, что второй подход полностью решает проблему редупликации. Таким образом, оба способа преодоления проблемы редупликации ошибочны. Остается искать другое решение проблемы тождества личности. Таким решением может служить биологический подход.

2. Биологический подход Э. Олсона

Некоторые философы полагают, что психологические характеристики принципиально не подходят для решения вопроса о тождестве личности. Такое представление построено на альтернативном взгляде на сущность и основу идентичности человека. Среди критиков психологического подхода – Э. Олсон, Д. Деграция, Дж. Томпсон, М. Эйрс, П. Сно-

удон, П. ван Инваген. Многие из этих философов замечают, что жизнь человека не сводится к наличию и сохранению психологических характеристик. Жизнь начинается с появления эмбриона, у которого нет развитых психических способностей, и часто заканчивается в ситуации, когда психика уже деградирует. В течение жизненного пути в результате травмы или болезни человек также может оказаться лишен каких-то когнитивных способностей: самосознания, рациональности, памяти. Однако это не является достаточным основанием вычеркивать соответствующие этапы жизни Я из автобиографии личности. Есть нечто более фундаментальное, базовое в характеристиках человека, всегда сопутствующее существованию его Я. Этой базой сторонники биологического подхода (или «анимализма») считают жизнь организма. В целях обоснования этой позиции один из самых ярких сторонников биологического подхода, Э. Олсон, разрабатывает *Аргумент вегетативного состояния* (АВС) и *Аргумент эмбриона* (АЭ).

А. Аргументы Э. Олсона в пользу биологического подхода

Допустим, пишет Олсон [Olson 1997, 7], что в результате несчастного случая человек оказывается в хроническом вегетативном состоянии. В результате перебоя работы сердца мозг десять минут не получает кислорода. Нейроны ко-

ры головного мозга гибнут от кислородного голодания, что приводит к перманентному отсутствию сознания. Речь, память и сложное поведение становятся принципиально невозможными. При этом сохраняются базовые функции: дыхание, пищеварение, кровообращение, метаболизм. Ведь субкортикальные части мозга, ответственные за эти функции (таламус, ствол мозга, мозжечок, базальные ганглии), более устойчивы к дефициту кислорода. Что можно сказать о таком состоянии?

Описанное состояние не является смертью, так как по всем важным признакам этот человек жив. Он может умереть, если перестать его кормить. Описанное состояние не является комой или сном. Человек в вегетативном состоянии переживает периоды сна и условного бодрствования. В этом состоянии человек даже реагирует на внешние стимулы: свет, звук. У него работают рефлексy. Он моргает и кашляет. Он даже может глотать пищу. Это парадоксальное состояние бодрствования без осознанности. Как говорят медсестры: свет горит, но никого нет дома. Но кто находится в этом состоянии? У кого горит свет? Может ли такое случиться с вами? Или в самом выражении «я буду находиться в вегетативном состоянии» уже есть противоречие? Будет ли такое состояние означать конец вашего существования или, несмотря на чудовищное качество жизни, это будет все еще продолжение жизни вашего Я?

Олсон склоняется к мнению, что даже такое существова-

ние – это продолжение жизни исходной личности, хотя и чудовищного качества. Чтобы склонить к этому заключению и свою аудиторию, он предлагает сравнить случай с жизнью в вегетативном состоянии со случаем, когда после смерти друзья делают памятник, точную копию человека из бронзы. Есть между этими случаями значимые отличия? – спрашивает он. Кажется, разница здесь принципиальная. В первом случае существование продолжается, а во втором – нет. Но на чем в таком случае основывается тождество личности человека в вегетативном состоянии и того, кем он был прежде? Возможны как минимум два варианта объяснений.

Во-первых, можно сказать, что между личностью до и после несчастного случая остается тождество в силу сохранения тела человека (большой его части), а во-вторых, что между ними существует тождество в силу сохранения преемственности биологических функций. Первый вариант ответа опирается на представление о том, что человек – это прежде всего тело, второй – что человек – это в первую очередь биологический организм, животное, *Homo sapiens*. Первый завязан на телесный критерий, а второй – на биологический. Между этими критериями существует определенное различие, но для целей данного исследования можно объединить эти два критерия в общий биологический подход⁴⁰. Исходя из этого, можно сформулировать новый общий критерий

⁴⁰ Такое объединение отчасти оправданно. На это указывает Олсон в тексте «Человеческое животное» [Olson 1997, 19].

тождества личности: если X – это личность в момент t_1 и Y существует в момент t_2 , X тождественен Y , если и только если биологический организм Y является преемником биологического организма X . Именно такой критерий я буду использовать для характеристики биологического подхода.

Вторым важным доводом в пользу биологического подхода является Аргумент эмбриона. Он опирается на очевидный факт: существование каждого индивида, каждой личности начинается с существования эмбриона. Но как связаны эмбрион и личность? Олсон считает их нумерически тождественными и предлагает доказательство этого от противоположного. Он допускает, что эмбрион и личность нумерически различны, но из этого следует одно из двух утверждений: (1) когда личность возникла, она заменила собой эмбрион в утробе матери; (2) с самого начала личность сосуществовала в одном пространстве с эмбрионом. Оба этих следствия Олсон считает абсурдными, и, следовательно, нужно сделать вывод о нумерическом тождестве эмбриона и личности. Это следствие, в свою очередь, доказывает, что психологическое тождество не необходимо для тождества личности, ведь у эмбриона нет психологических характеристик. Таким образом обосновывается анимализм.

Биологический подход, в частности подход Олсона, предполагает согласие с двумя утверждениями. Во-первых, что «мы с вами являемся животными: представителями вида *Homo sapiens*, если быть точным» [Olson 1997, 17]. Во-вто-

рых, что условием продолжения нашего существования является продолжение существования животного организма, а не продолжение существования сознания, психики или чего-то еще. Первое убеждение обычно не вызывает сомнения, в особенности потому, что оно не исключает возможности существования других типов личностей. «Насколько я знаю, в мире существуют или могли бы существовать разумные марсиане, боги, ангелы, демоны, тролли и даже рациональные, сознательные электронные компьютеры, сделанные из металла и кремниевых микросхем» [Olson 1997, 17]. Просто это не наш случай, считает философ, мы – биологические существа. Но между этим очевидным утверждением и вторым, менее очевидным, Олсон прослеживает логическую связь.

Если мы – биологические существа, то именно в продолжении жизни биологического существа будет заключаться наше длительное существование. Из этого, в свою очередь, следует, что преемственность психологических характеристик не является необходимым условием тождества личности. «Очевидно, не всем человеческим животным требуется психологическая преемственность для продолжения существования. Например, люди, находящиеся в вегетативном состоянии, человеческие эмбрионы, анэнцефалитики – это человеческие животные, которые выживают без каких-либо психологических свойств» [Olson 1997, 18].

В. Сравнение биологического, физиологического и психологического подходов

По мнению Олсона, биологический подход следует отличать от физиологического подхода. Согласно физиологическому подходу, продолжение нашего существования основано на сохранении большей части мозга. Мозг выделяется как основной орган в первую очередь потому, что он является носителем психики. Именно поэтому физиологический подход правильней всего относить к разновидностям психологического. Это важное замечание, так как здесь термины способствуют путанице. Не менее важно проводить различие между биологическим подходом Олсона (чья позиция наиболее типична для сторонников данной стратегии) и теорией Д. Виггинса. Последнего тоже иногда называют теоретиком «биологического подхода», однако даже в поздних работах Виггинс утверждает важность некоторых психологических черт для сохранения тождественности личности. Различия психологического и биологического подходов лучше всего проследить на гипотетических примерах.

Первый из них – уже упомянутый выше «Телетранспортер» Д. Парфита. Напомню, в этом сценарии предлагается перемещение личности с помощью передачи данных о ее физическом носителе, организме. По условиям мысленного

эксперимента Парфита, исходный организм уничтожается, а в другом месте «собирается» новый. Для сторонника психологического подхода «Телетранспортер» Парфита является средством перемещения. Но для сторонника биологического подхода это – камера смерти. Анималисты уверены, что, поскольку биологическая жизнь исходной личности прерывается, личность также уничтожается и нумерически. И это является концом жизни индивида. Различие этих двух подходов можно продемонстрировать также на примере другой гипотетической ситуации.

В мысленном эксперименте «Трансплантация мозга» из тела одного человека забирается головной мозг и пересаживается в тело другого человека. Трансплантация проходит успешно, и человек полностью реабилитируется. У него восстанавливаются все когнитивные функции, в том числе воспоминания, характер и т. п. Хотя, конечно, некоторое время ему непривычно в его новом теле. Как интерпретировать данную ситуацию? Для сторонника психологического подхода это способ обретения нового тела, то есть трансплантация тела. А для сторонника биологического подхода это выживание без мозга, или донорство мозга. Как видно, подходы ведут к противоположным суждениям. Но достаточно ли обоснован биологический критерий тождества личности?

С. Критика биологического подхода

Биологический подход справляется с отождествлением личности в различных сложных ситуациях: когда человек спит, когда находится в вегетативном состоянии или только появляется на свет. Помимо этого, биологический подход, кажется, наиболее близок бытовой практике идентификации. Ведь именно на внешние признаки, на продолжение существования организма мы чаще всего ориентируемся, когда устанавливаем идентичность личности. И тем не менее биологический подход, как и психологический, обладает существенными недостатками. Основным недостатком я считаю саму постановку проблемы «тождества личности» в анимализме.

Выше я определял проблему асинхронного тождества личности как проблему реидентификации личности, то есть как вопрос о критерии тождества личности в момент t_1 с личностью в момент t_2 . Решение проблемы в этой формулировке имеет практическое значение. В частности, оно нужно для приписывания моральной ответственности и выявления условий выживания. Олсон меняет эту формулировку. Он предлагает искать критерии тождества не между личностями, а между личностью и, например, объектом или животным. Само по себе это интересное занятие. Но, с моей точки зрения, это новое направление поиска дает нереле-

вантные результаты. Анималистическая реидентификация не позволяет решить практические задачи приписывания ответственности и определения случаев выживания. Для иллюстрации предложу гипотетический случай «Омар».

Омар. Некоторое время назад жил человек по имени Грегор Замза. Грегор участвовал в спасательных работах на атомной электростанции, в результате чего подвергся сильному радиационному облучению. После облучения его клетки стали мутировать, и за несколько лет Грегор превратился в Омара. Все основные биологические функции его организма сохранились, однако он потерял большую часть своих психических способностей. После метаморфоз его психика стала полностью соответствовать психике ракообразного. Через некоторое время Омар еще раз подвергся сильному радиационному облучению и стал вновь меняться. Результатом метаморфоз стало появление нового существа по имени Кейти. Кейти – человек женского рода, и она обладает только смутными воспоминаниями о ранних стадиях метаморфоз. Все биологические функции Кейти преемственны с Омаром, но ее психика полностью отличается как от его психики, так и от психики Грегора.

Согласно биологическому подходу, в описанной выше гипотетической ситуации Грегор тождественен Омару, Омар – Кейти и Кейти – Грегору. Но что означает эта тождественность? Можно ли считать, что Грегор выжил как Омар или Кейти? Или что Кейти несет ответственность за поступки

Грегора? Кажется, нет. Несмотря на сохранение биологических функций, между Грегором и Кейти нет достаточной связи, чтобы транслировать ответственность. Значит, отношение биологического тождества не решает практические вопросы. Даже если Грегор как организм продолжает жить в качестве Кейти, Кейти не несет ответственности за его действия. Но тогда биологическая тождественность не решает нашу проблему. Возможно, для каких-то целей реидентификация с помощью биологического подхода полезна, но в контексте нашего исследования она не подходит. Значит, нам нужно искать другой метод. Такой альтернативой является субстанциональная теория личности.

3. Субстанциональный подход и его сторонники

Согласно субстанциональному подходу, тождество личности заключается в тождестве субстанции. Иными словами, чтобы личность P_2 в момент t_2 была той же самой, что личность P_1 в момент t_1 , у них должна быть нумерически одна субстанция. Под субстанцией обычно подразумевается принципиально единая и неделимая сущность вроде живой нематериальной души, ментальной вещи, способной существовать автономно от тела и мозга. То есть сторонник этой теории должен занимать дуалистическую позицию в отношении проблемы сознание-тело. Субстанциональный под-

ход является, по определению Д. Парфита, ведущего современного теоретика тождества личности, «простым решением» проблемы тождества в том смысле, что он предполагает одно-единственное обстоятельство, определяющее тождественность, – неизменную духовную монаду. Простой подход не сводит тождество личности к тождеству или преемственности каких-то других свойств. В этом смысле он отличается от описанных выше «комплексных решений»: психологического и физиологического, которые по своей сути являются редуционистскими.

Субстанциональный подход обеспечивает психологический комфорт, создавая личности как бы дополнительный контур безопасности. Возможно, в силу этого он разделяется большинством неспециалистов. В рамках этой теории можно непротиворечиво помыслить реинкарнацию, воскрешение в теле или бестелесную жизнь после смерти, переселение в другое тело или даже совместное использование двумя и более ментальными субстанциями одного тела. То есть эта теория согласуется с распространенными представлениями о возможностях существования и выживания личности. Субстанциональный подход позволяет также решить вопрос о единстве сознания, об уникальной природе «личности» (через нумерическую уникальность субстанции), установить тождество между бессознательными фазами жизни человека (в том числе в состоянии эмбриона и вегетативном состоянии) и зрелой сознательной личностью, объяс-

нить особую заботу личности о своем будущем и моральную ответственность. Возможно, с учетом тезиса о неделимости ментальной субстанции он также позволяет разрешить описанные выше парадоксы редупликации личности. В силу множества этих обстоятельств его разделяют и некоторые философы.

Самым знаменитым сторонником «простого» решения был Р. Декарт. Вслед за Декартом «простое» решение проблемы тождества личности принимали также Дж. Батлер и Т. Рид. Среди современников наиболее известными философами, отстаивающими эту позицию, считаются американец Р. Чизом [Chisholm 1991] и британец Р. Суинбёрн. Здесь я остановлюсь подробнее на аргументах оксфордского философа [Swinburne 1991; Swinburne 2013]. Его защита субстанциональных критериев тождества построена на современной критике психологического и физиологического подходов и наиболее любопытна.

А. Субстанциональный поход Р. Суинбёрна

Суинбёрн различает эпистемологический и метафизический уровни решения проблемы тождества личности. Первый уровень, по его мнению, относится к практической задаче реидентификации, то есть заставляет исследователей искать наглядные свидетельства тождественности. Второй уровень относится к подлинному основанию тождественности.

Философ считает, что «эмпирицистские теории» – так Суинбёрн называет психологический и физиологический подходы (отметим, что он не выделяет «биологический» подход отдельно, но его аргументы против физиологического подхода могут быть отчасти использованы и против биологического) – успешно справляются с первой задачей, но неспособны решить вторую. Память, а также психологическая и физиологическая преемственность действительно являются надежными сопутствующими критериями, но эти факторы не являются основой тождественности личности.

Ключевым основанием для этого вывода является неизбежное, по мнению Суинбёрна, следствие эмпирицистских теорий: в некоторых гипотетических случаях, согласно этим теориям, не существует правильного ответа на вопрос, являются ли две личности тождественными. Он убежден, что все эмпирицистские теории ведут к абсурдным следствиям. Доказательство этого начинается с новой классификации, которую предлагает оксфордский философ и которая, по его замыслу, исчерпывает эмпирицистские альтернативы. Эта классификация предполагает три типа теорий тождества личности. Против каждого типа Суинбёрн использует отдельный аргумент. Результатом этих рассуждений должно стать опровержение всех эмпирицистских теорий.

В. Недостатки эмпирицистских теорий

Первый тип эмпирицистских теорий предполагает возможность редупликации. Такими теориями являются классические теории, основанные на психологических критериях: памяти и характере. Согласно таким теориям, личность P_2 в момент t_2 является тождественной личности P_1 в момент t_1 , если удовлетворяет какому-то критерию, при этом более одной личности может удовлетворять этому критерию. По мнению Суинбёрна, такие теории являются ложными потому, что неизбежным следствием их является противоречие. Одна личность в прошлом, согласно этим теориям, может оказаться двумя и более личностями в будущем, как в случае с Гаем Фоксом, Чарлзом и Робертом. Но две личности не могут быть тождественны одной, а также две тождественные личности не могут быть в разных местах. Следовательно, критерий тождества, выдвигаемый этими теориями, неверен.

Теории второго и третьего типа не допускают редупликации, но это достигается разными способами. Теории второго типа – это теории первого типа, к которым просто добавлено дополнительное условие, запрещающее редупликацию. Это условие может звучать так: личность в момент t_2 не может быть той же личностью в момент t_1 , если в момент t_2 существует еще какая-то личность, удовлетворяющая тем

же критериям тождества. Не только физиологические, но и психологические теории в целом могут относиться ко второму типу. Например, ко второму типу относятся теории, согласно которым утверждается, что P_2 в момент t_2 – это та же личность, что P_1 в момент t_1 , если и только если существует пространственно-временная преемственность между какой-то частью мозга P_1 и P_2 и не существует аналогичной преемственности между какой-то частью мозга P_1 и другой личностью, кроме P_2 .

Теории второго типа, по мнению Суинбёрна, тоже ведут к абсурдному следствию. Этим следствием является то, что решение о тождестве и выживании некоторой личности в этих теориях зависит от судьбы каких-то других личностей. Примером может служить случай с трансплантацией двух полушарий разным донорам. Согласно теориям второго типа, выживание личности будет возможно, только если один из доноров погибнет, а это крайне контринтуитивное следствие. Таким образом, заключает Суинбёрн, теории этого типа тоже ошибочны.

Теории третьего типа также не допускают редупликацию, но в них нет дополнительного условия. Просто в этих теориях критерии тождества определены так, чтобы им могла соответствовать только одна личность в будущем. Примером может служить теория, которая основывается на следующем физиологическом критерии тождества: P_2 в момент t_2 – это та

же личность, что P_1 в момент t_1 , если и только если существует такая пространственно-временная преемственность между организмом P_1 и P_2 , что на каждой последующей временной стадии между t_1 и t_2 организм имел более 90 % мозга тождественными соответствующей части мозга этого организма на предшествующей стадии. Такое условие не допускает «разделения» организма, ибо тогда не будет удовлетворено условие наличия 90 % предшествующего мозга.

Теории такого типа тоже ведут к абсурдному следствию. Этим абсурдным следствием является то, что одна и та же личность может являться и не являться тождественной другой личности в зависимости от того, каким образом она перенесла изменения. Так, личность, единомоментно потерявшая 50 % мозга, не выживет (так как она сохранит менее 90 % мозга), но если она потеряет 50 % мозга в результате череды потерь, каждая из которых будет составлять менее 10 %, то она продолжит существование.

Кроме того, критерий тождества, предлагаемый такой теорией, представляется произвольным. Почему 90 %? Почему 70 % или 80 % мозга недостаточно для выживания личности? Как решить, где именно должна проходить грань, обеспечивающая тождество? Какое бы процентное отношение ни было поставлено, оно будет выглядеть искусственным и произвольным. Кроме того, всегда остается возможность для пограничного случая. Предположим, теория допускает, что для выживания достаточно сохранения не менее 50 %

мозга. Тогда представим случай, что именно 50 % мозга сохраняется у пациента после вынужденной операции. Прекратится ли его существование (даже если пациент выживет и сохранит когнитивные способности)? Ответ не очевиден. Это, по мнению оксфордского философа, свидетельствует о ложности теорий третьего типа.

Суинбёрн считает, что фундаментальная проблема этих теорий в том, что они вопрос о тождестве личности уподобляют вопросу о тождестве физических объектов. «Нет ничего загадочного в отношении пограничного случая, когда новый автомобиль сконструирован из частей моей старой машины и новых частей. Поскольку у машин нет чувств и надежды на выживание... Но сознающий объект, которым является личность, желает оставаться в сознании, и кажется, что не может существовать промежуточной возможности между тем, чтобы это была та же личность или другая», – считает он. По мнению оксфордского философа, личности (как объекты с сознанием) принципиально отличаются от физических сущностей и имеют особые условия тождественности.

С. Особенности субстанционального подхода суинбёрна

Сторонники эмпирицистских теорий не проводят различия между тождеством физических объектов и тождеством

личности, и Суинбёрн считает это основанием для сомнения в их истинности. Единственным решением этой метафизической головоломки, с его точки зрения, может быть субстанциональный подход. В основании его субстанционального подхода – несколько утверждений, которые я обозначу как тезисы о (1) скрытом основании, (2) определенности и (3) дискретности отношения тождества личности. Совокупность этих тезисов составляет содержательное ядро позиции Суинбёрна. Согласно идее о нередуцируемости, тождество не заключается исключительно в преимственности одной или нескольких наблюдаемых характеристик... Единственная альтернатива – сказать, что тождество личности – это нечто фундаментальное. То есть, согласно теории Суинбёрна, единство личности несводимо к каким-то другим отношениям.

Фундаментальным, нередуцируемым отношением, по мнению философа, является тождество субстанции, нематериальной сущности. Это утверждение я называю *Тезисом о скрытом основании*. Основным свойством этой сущности является способность «иметь сознательные состояния (например, мысли и ощущения) и осуществлять намеренные действия» [Swinburne 1991, 27]. Согласно Тезису о скрытом основании, отношения между наблюдаемыми характеристиками представляют лишь косвенные свидетельства тождества субстанции, но не гарантируют его.

Одна нематериальная субстанция всегда либо тожде-

ственна, либо не тождественна другой субстанции в другое время. Это – второй тезис, который я называю *Тезисом об определенности*. В эпистемологическом плане даже наличия полной информации о наблюдаемых характеристиках, этих косвенных свидетельствах тождества, может быть недостаточно для того, чтобы сделать вывод о наличии или отсутствии тождества личности со всей определенностью. Но это не означает, что одна личность может стоять в некотором неопределенно-тождественном отношении к другой личности в другой момент времени. Второй тезис утверждает, что это отношение всегда определено.

Эмпирицистские теории склоняют нас к тому, чтобы считать тождество вопросом степени. То есть они допускают истинность высказываний «он в большей степени тот же человек, что и был». Но, по мнению Суинбёрна, эти высказывания ложны. Высказывания о тождестве личности всегда дискретные: личность может быть либо тождественна другой, либо отлична от нее. И промежуточных случаев быть не может – личность не может быть в какой-то степени тождественной. Это утверждение я буду называть *Тезисом о дискретности*. Можно ввести понятие тождества личности со степенью тождества – по аналогии с тождеством физических объектов, которое может рассматриваться как недискретное. Но этот ход непродуктивный, так как он, с позиции Суинбёрна, не относится к понятию тождества личности, с которым мы имеем дело в обыденном языке. Все три тезиса, на-

против, соответствуют обыденным интуициям в отношении тождества личности.

Д. Критика субстанционального подхода с помощью верификационизма

Субстанциональный подход имеет ряд описанных выше преимуществ: он согласуется с обыденными интуициями в отношении тождества и реидентификации личности, решает запутанные ситуации с редупликацией и разделением личностей и т. д. Но он сам является мишенью для критики. Против него можно использовать аргументы, построенные на Принципе верификации. Этот принцип защищался логическими позитивистами в первой половине XX в. Однако в его основе лежат идеи, заложенные еще Д. Юмом. Как пишет историк философии, исследователь верификационизма Ш. Мисак: «За руководящую мысль они (логические позитивисты. – Д. В.) взяли исходную идею Юма о том, что наши убеждения, если они вообще обоснованы, обоснованы данными опыта» [Misak 2005, 56]. Согласно одной из формулировок Принципа верификации, «осмысленное предложение – это такое предложение, истинность или ложность которого в конечном итоге может быть подтверждена на практике» [Misak 2005, 62].

Опираясь на Принцип верификации в такой формулировке, можно попытаться доказать, что утверждение о субстан-

циональной основе тождества личности не является осмысленным. Для этого необходимо применить этот принцип к Тезису о скрытом основании тождества из теории Суинбёрна. Согласно Тезису о скрытом основании тождества, наблюдаемые характеристики могут лишь косвенным образом способствовать реидентификации. То есть любые внешние свойства, доступные в опыте, не являются достаточным основанием для реидентификации. Таким образом, согласно субстанциональному подходу, на практике невозможно окончательно подтвердить тождество личности в разное время. Но это противоречит Принципу верификации в описанной выше формулировке. Значит ли это, что субстанциональный подход выдвигает бессмысленный критерий тождества личности?

Такой верификационистский аргумент мог бы быть использован против подхода Суинбёрна. Однако посылка в этом аргументе – не самая надежная. Принцип верификации в такой формулировке, в его сильной версии, сам уязвим для критики. На это указывает Суинбёрн: «Что касается меня, я не нахожу оснований принимать любые из этих видов верификационизма. До тех пор пока мы понимаем грамматическую форму повествовательного предложения и слова, которые в нем употребляются, мы понимаем, что в нем утверждается, даже если нет мыслимых способов подтвердить его истинность или ложность» [Swinburne 1991, 39–40]. По мнению оксфордского философа, сильная версия верификаци-

онизма предъявляет слишком высокие требования, так что даже тривиальные утверждения попадают в категорию бессмысленных. Это сводит Принцип к абсурду. В качестве примера он приводит выражение «в комнате стоит стол». Это утверждение нельзя абсолютно точно подтвердить или опровергнуть, так как восприятие стола может быть иллюзией или сном. Значит, согласно Принципу верификации в сильной версии, это бессмысленное выражение. Но это абсурд, и, значит, Принцип верификации в его сильной формулировке ложен.

Суинбёрн считает, что для осмысленности выражения достаточно, чтобы был *хоть какой-нибудь способ вероятностной верификации или фальсификации высказывания*. Это утверждение можно считать слабой версией верификационизма. Оно вполне правдоподобно. Философ предлагает рассмотреть три предложения: «на нашей планете нет единорогов»; «в нашей Галактике нет единорогов»; «в нашей Вселенной нет единорогов». Практическая возможность проверки всех трех утверждений разная, но представляется, что они в равной степени имеют смысл. Из этого, по мнению Суинбёрна, следует, что любой способ вероятностной верификации или фальсификации достаточен для осмысленности высказывания. «До тех пор пока мы требуем только возможности подтверждения или опровержения, будь то возможность логическая, физическая или практическая, доступной хотя бы одной личности или кому-нибудь, сейчас или ко-

гда-нибудь, дуалистическая теория тождества личности не будет уязвима для верификационизма» [Swinburne 1991, 40–41]. Таким образом, Суинбёрн опровергает сильную версию и соглашается со слабой версией верификационизма. При этом он считает слабую версию верификационизма совместимой с утверждением о субстанциональных критериях тождества личности.

Действительно, психологические и физические характеристики не дают безусловной гарантии тождества. Но, согласно субстанциональному подходу Суинбёрна, наличие преемственности психологических, физиологических и биологических характеристик позволяет косвенным образом с некоторой вероятностью реидентифицировать личность. Так, сохранение большей части мозга и воспоминаний в рамках субстанционального подхода является подтверждением, что личность P_2 – это та же личность, что и личность P_1 . А появление конкурента P_2^* с большей частью мозга и воспоминаний P_1 – фальсификацией отождествления P_2 и P_1 . То есть, по мнению Суинбёрна, существуют вероятностные способы верификации субстанциональных критериев тождества личности, а значит, они осмысленны.

Ответ оксфордского философа на критику с помощью верификационизма я считаю не до конца удовлетворительным. Вопреки утверждениям Суинбёрна, я считаю, что даже слабая версия верификационизма несовместима с субстанциональным подходом к тождеству личности. Оксфордский фи-

лософ утверждает, что биологические, физиологические и психологические характеристики могут служить косвенными основаниями для высказываний о тождестве. Но эти характеристики могли бы быть косвенными свидетельствами тождества субстанции, только если бы они были должным образом связаны с субстанцией. Эта особая связь могла бы быть отношением тождества, супервентности или каузальной зависимости. Но таких связей, как мне представляется, нет. Отношение тождества между нематериальной субстанцией и психологическими, физиологическими или биологическими характеристиками должно быть отвергнуто, так как, по определению Суинбёрна, нематериальная монада обладает совсем иными свойствами: она неделима и ее сущностные свойства не могут быть скопированы. Отношение супервентности предполагает прямую зависимость свойств: набор свойств B является супервентным по отношению к набору свойств A в том случае, если объект не может различаться в свойствах B в отсутствие различий в свойствах A . Из теории Суинбёрна следует, что субстанция может быть неизменной при изменении указанных характеристик. Следовательно, между свойствами субстанции и этими свойствами не существует отношения супервентности. И наконец, тезис о каузальной связи между субстанцией и психологическими и телесными свойствами крайне проблематичен.

Очень мало философов допускают возможность участия субстанций в каузальных связях. Наиболее распространен-

но мнение о том, что только события могут находиться в каузальной связи. Более того, каузальная эффективность нематериальной субстанции крайне экзотична и плохо согласуется с правдоподобным тезисом о каузальной замкнутости физического мира. Нужно опровергнуть много аргументов, чтобы доказать, что нематериальная субстанция может быть каузально связана с психическими, физиологическими и биологическими характеристиками. Таким образом, связь между монадой и этими свойствами, скорее всего, аналогична связи между изображениями единорогов, описаниями единорогов, продаваемыми «рогами» единорогов и единорогами. Эти «косвенные доказательства» существования единорогов не подтверждают их существования потому, что между ними и единорогами отсутствует должная связь. Эти «доказательства» не тождественны единорогам, не супервентны и каузально не связаны с существованием единорогов. Описания этих удивительных животных не связаны каузально с восприятиями единорогов, а «рога» на самом деле являются рогами нарвалов. Как из обнаружения большего числа текстов или «псевдорогов» не следует увеличение вероятности существования единорогов, так и из наличия большего числа преемственных психологических и прочих характеристик не следует большая вероятность тождества монад. Отсутствие возможности хотя бы косвенным образом удостовериться в тождественности или различии душ в разные моменты времени, в соответствии с верификаци-

онизмом, говорит о том, что субстанциональный критерий тождества личности является бессмысленным.

III. Критика классических подходов с помощью Аргумента пяти трансплантаций

Похоже, мы оказались в патовой ситуации. Субстанциональный подход к тождеству личности неудовлетворителен из-за отсутствия способа верификации. Психологические теории уязвимы перед Аргументом редупликации, а биологические – подменяют исходную проблему. Цепочку рассуждений, приведшую нас к этому пату, можно резюмировать с помощью последовательности мысленных экспериментов с «трансплантациями». Пять трансплантаций позволяют оценить гипотезы об основаниях и критериях тождества личности, предложенные современными философами, и выявить их ключевые недостатки.

A. Трансплантация души

Гипотеза 1. Критерием тождества, необходимого для моральной ответственности, является тождество души.

Предположим, Браун и Робинсон обладают душами вдобавок к телам и их физиологическим, биологическим и психологическим характеристикам. У каждой из личностей есть

свое уникальное содержание психики и различные психические способности. В какой-то момент их души поменялись местами: душа Брауна трансплантируется в тело Робинсона, и наоборот. При этом, поскольку психологические характеристики супервентны на физических, все их психические свойства остались неизменными: они продолжают иметь те же воспоминания, те же намерения и желания, те же психические качества и характер. Как можно будет определить, что Браун и Робинсон поменялись душами?

Очевидно, этого нельзя будет сделать со стороны: с точки зрения внешнего наблюдателя, никакой перемены не произошло. Браун и Робинсон не изменились в поведении. Но и изнутри, в интроспекции Браун и Робинсон не смогут обнаружить этой перемены, ведь у них сохранятся все те же воспоминания, привычки и прочие психологические особенности, что и были. Так стал ли Браун Робинсоном, а Робинсон – Брауном? Кажется, что этого не произошло. Должен ли теперь Робинсон нести ответственность за поступки, совершенные Брауном? Скорее всего, нет. Каждая личность осталась на своем месте. Это позволяет думать, что субстанциональный критерий тождества личности не согласуется с нашими интуициями. Тогда вместо субстанционального подхода испытаем биологический подход.

В. Трансплантация мозга

Гипотеза 2. Критерием тождества, необходимого для моральной ответственности, является тождество/преемственность биологических функций.

Представим, что злодей-нейрохирург провел эксперимент: пока Браун и Робинсон спали, он перенес мозг Брауна в тело Робинсона, а мозг Робинсона – в тело Брауна. Проснувшись, тело Брауна стало вести себя совсем как Робинсон, а тело Робинсона – как Браун: тело Робинсона теперь демонстрирует наличие характера, воспоминаний и привычек, что были у Брауна, и наоборот. И теперь обе личности, Браун и Робинсон, не узнают себя в зеркале. Где Браун, а где – Робинсон? Согласно биологическому подходу, мы должны сказать, что личности сохранили свое местоположение. Но интуиции подсказывают иное.

Кажется, что в результате трансплантаций мозгов личности переместились, хотя тела и прочие органы остались на своих местах. Оправданно ли теперь будет применять ответственность (например, наказание) за поступки Брауна к телу Брауна? Кажется, нет. Похоже, что биологический критерий приводит к контринтуитивным выводам и поэтому ложен. Возможно, основой тождества личности является тождество мозга?

С. Трансплантация психики

Гипотеза 3. Критерием тождества, необходимого для моральной ответственности, является тождество мозга.

Предположим, ученые научились копировать информацию, которая хранится в нейронных сетях мозга. Эта информация представляет все содержание психики и все психические способности. Команда ученых провела эксперимент: они скопировали все данные из мозга Робинсона в мозг Брауна, и наоборот. Мозги Брауна и Робинсона остаются на месте, но все их психические особенности были перенесены. Теперь Робинсон ведет себя так, как раньше вел себя Браун. Все воспоминания Брауна стали воспоминаниями Робинсона. И наоборот. Где Браун, а где Робинсон?

Интуиции склоняют к тому, чтобы признать смену тел у личностей. Биологические органы остались на месте, мозги остались на месте, но личности мигрировали. Но это противоречит представленной гипотезе. Выходит, что тождество мозга не является фундаментальной основой тождества личности. Возможно, носителем личности является психика, а не физический орган. Это дает возможность надеяться на продолжение жизни личности после смерти организма. Если, конечно, ученые научатся копировать всю информацию о психике на новый носитель. Но для оптимизма пока недо-

статочно оснований. Стоит проверить и эту гипотезу.

Д. Трансплантация психики с умножением личности

Гипотеза 4. Критерием тождества, необходимого для моральной ответственности, является тождество/преемственность психики.

Предположим, что после копирования информации из мозга Робинсона в мозг Брауна ученые скопировали эту информацию еще и в мозг Джонсона. После этой процедуры Браун и Джонсон приобретают характер, привычки, намерения, желания и воспоминания Робинсона. Когда их спрашивают имя, они хором отвечают – Робинсон. Когда их спрашивают о биографии, они рассказывают биографию Робинсона. Браун и Джонсон – полные психологические копии Робинсона. При этом они – разные люди. Являются ли они одной и той же личностью? Несут ли они ответственность за поступки, совершенные Робинсоном в прошлом?

Кажется, что здесь интуиции начинают подводить. Этот случай настолько фантастический, что полагаться на опыт становится невозможно. Сходных ситуаций в жизни не встретишь. Что-то подсказывает, что оба этих человека – Робинсон. Но как Робинсонов может быть много? Нет ли здесь противоречия? Как может быть множество идентичных людей? Кто теперь женат на миссис Робинсон? Кто имеет пра-

во получать пенсию Робинсона и отвечать за его поступки? Эти вопросы могут поставить в тупик. Судя по всему, возможность трансплантации психики сразу нескольким агентам ставит под сомнение и психологический подход к тождеству личности. Но если вариант умножения личности кажется малопонятным или недостаточно убедительным, вместо него против психологического подхода можно использовать мысленный эксперимент с делением личности. Его эмпирическая возможность оговаривалась выше.

Е. Трансплантация психики с делением личности (Fission case)

Носителем высших психических способностей является мозг. Кажется, поэтому у нас есть интуиция, что при трансплантации мозга происходит трансплантация личности. Мозг, как известно, состоит из двух полушарий. При этом каждое полушарие автономно, может изолированно выполнять многие высшие психические функции. Существует способ лечения обширной эпилепсии путем удаления одного из полушарий мозга – гемисферэктомия. Если операция проводится в детском возрасте, благодаря нейропластичности все когнитивные функции восстанавливаются.

Предположим, что Вильямсон страдает редким заболеванием, которое связано с нарушением взаимодействия двух полушарий мозга. В качестве лечения ему предлагают опе-

рацию: пересадку обоих полушарий в организмы доноров – Робинсона и Брауна. Мозги доноров полностью дисфункциональны, Робинсон и Браун находятся в хроническом вегетативном состоянии. Допустим, операция проходит удачно, и мозги Вильямсона приживаются в организмах Робинсона и Брауна. Тело Вильямсона остается в вегетативном состоянии. Где теперь Вильямсон? Тела Робинсона и Брауна теперь имеют психические способности Вильямсона, его убеждения и желания. Кто из них тождественен Вильямсону?

Предположим, Робинсон позавтракал, а Браун не успел, тогда Робинсон сытый, а Браун – голодный. Если обе эти личности – Вильямсон, то Вильямсон и сытый, и голодный. Это противоречие. Значит, этот пример ставит под сомнение психологический критерий. Дело в том, что *тождество – это отношение один к одному, а психологический подход допускает отношение один ко многим*. Или Вильямсон уже изначально был множественен и мы просто могли об этом не знать?

Случаи с умножением или делением личности являются проблемой для психологического подхода. Но в совокупности представленные гипотетические ситуации ставят под сомнение все описанные выше подходы к проблеме тождества личности. Аргумент пяти трансплантаций демонстрирует проблему и является серьезным препятствием для признания возможности свободы и моральной ответственности. Ключ к разрешению патовой ситуации можно попытаться

отыскать в другом мысленном эксперименте, фантастической истории, которая подводит нас к новой концепции личности. История эта называется «Где Я?», а ее автор – уже известный нам Д. Деннет.

IV. Теория личности Д. Деннета

1. Мысленный эксперимент «Где Я?»

История «Где Я?» появилась впервые в сборнике «Мозговые штурмы» (1978). Она была включена в завершающую главу, которая значительно отличалась от всех предыдущих. Автор во введении сам назвал ее «десертом» к трудным, серьезным философским рассуждениям. Повторно эта история была опубликована в книге «Глаз разума. Фантазии и размышления о самосознании и о душе», вышедшей под совместной редакцией Д. Хофштадтера и Д. Деннета. По своему жанру история «Где Я?» гораздо естественней подходила именно для этой публикации, в которой собрано, пожалуй, только «сладкое»: эссе философов, писателей, ученых (Р. Нозика, Х. Борхеса, С. Лема, Р. Докинза, Дж. Сёрла), размышляющих над вопросами «Что такое разум?», «Что такое Я?», «Может ли материя думать и чувствовать?», дополненные комментариями редакторов. История «Где Я?» по своему стилю не похожа на традиционные для аналитической философии мысленные эксперименты: выводы, следующие из этой истории, сразу не очевидны. Это – «попытка выявить возникающие... противоречия и описать их живо и образно. Мы хотели не столько ответить на вопросы, сколько взбудо-

ражить всех, кто прочтет нашу книгу...» [Хофштадтер, Деннетт 2003, 1], – признаются Деннет с Хофштадтером. История «Где Я?» настолько художественна, что даже стала одним из сюжетов одноименного фильма голландского режиссера Пьета Ходедреса, где в роли *Деннета* снимается сам Деннет. Вместе с ним героями фильма оказались его соредатор, философ-математик Дуглас Хофштадтер, и когнитивист Марвин Мински.

Как и прочие эссе сборника «Глаз разума», фантастический рассказ Деннета оставляет читателя в задумчивости. Он посвящен вопросу тождества личности и самоидентичности. Но какой подход он представляет? В некотором смысле история «Где Я?» похожа на уже упомянутый мысленный эксперимент с двумя черными ящиками, приведенный в первой главе. Здесь автор также не разъясняет свои выводы явным образом. Вместо этого в тексте мы находим разные, порой трудно соотносимые размышления героя истории и фрагментарные замечания автора. Какова позиция Деннета? Как и рассказ о хитроумной диковинной археологической находке, история «Где Я?» заставляет задуматься: «Как можно ее приспособить?», «Для чего бы она могла пригодиться?», «Каков был замысел создателя?», «Сделана ли она всерьез или для забавы?». В надежде прояснить альтернативы решения проблемы тождества личности мы проведем исследование этого фантастического сюжета.

Итак, *Деннет* выполняет секретную миссию Пентагона

(конечно же, против «красных»). Ему придется обезвредить радиоактивную боевую головку где-то в недрах земли. Но для того, чтобы мозг не был поврежден смертоносным излучением (для других органов излучение относительно безопасно), он должен остаться в лаборатории в Хьюстоне. Оттуда мозг будет осуществлять контроль над телом с помощью системы радиосвязи. Каждое нервное окончание в пустом черепе будет соединено с микропередатчиками и микроприемниками, которые, в свою очередь, будут направлять сигнал непосредственно в мозг. «Представьте, что это всего лишь *растяжение* нервов... Мы всего лишь сделаем ваши нервы бесконечно растяжимыми, вставив в них радиоконтакты» [Хофштадтер, Деннетт 2003, 192]. *Деннет* кладут на операционный стол, проводят анестезию, вынимают мозг из черепной коробки и помещают его в чан с живительным раствором. И вот, сразу после пробуждения, Деннет задает обычный для такого случая вопрос: «Где я?» Экстравагантная завязка нужна автору для того, чтобы обнаружить проблематичность этого вопроса. В самом деле, где *Деннет*, когда части его – в разных местах: мозг – в чане, а тело – на операционном столе?

Конфигурация 1. Гамлет (тело) и Йорик (мозг)

Подобного рода вопрос уже возникал раньше, когда об-

суждались различные варианты трансплантаций. Первая конфигурация Деннета отчасти близка случаю с трансплантацией мозга. Но в той ситуации ответ казался очевидным: личность была там, где мозг. Когда производилась трансплантация мозга, вместе с ней трансплантировалась и психика личности. Это, скорее всего, предопределяло работу интуиции. В данном же случае все не так однозначно. *Мозг Деннета трансплантирован, но тело сохранило функциональность и свою связь с мозгом и психологическими характеристиками.* Это запутанный вариант трансплантации. «Итак, я сижу на складном стуле и смотрю сквозь небьющееся стекло на свой собственный мозг. Но погодите... не должен ли я сказать, что плаваю в булькающей жидкости, озираемый собственными глазами?» [Ibid., 193]. Герой мысленного эксперимента, глядя на собственный мозг, философствует в шекспировском духе. И автор сюжета решает не ограничиваться намеком на знаменитую трагедию. «Совсем сбитый с толку, я попытался прибегнуть к любимому трюку философов. Я принялся раздавать вещам имена. “Йорик, – сказал я вслух своему мозгу, – ты – мой мозг. Остальное тело, сидящее на этом стуле, я назову “Гамлетом””. Все мы сейчас здесь: мой мозг, Йорик, мое тело, Гамлет[,] и я сам, Деннетт. Где же я *теперь?*» [Ibid., 193].

Альтернатива 1. Где Йорик, там и Деннет. Такое решение подходит некоторым сторонникам психологического

подхода. В первую очередь – тем из них, кто ассоциируют тождество личности именно с мозгом⁴¹. В мозге запечатлены подробности жизни, ключевые психологические качества и способности личности. Но, как я отметил, случай трансплантации в истории Деннета особенный. Несмотря на то что «управление» организмом ведется из «чана» в научной лаборатории, перспектива героя – со стороны его тела. И основные жизненно важные события для *Деннета*, кажется, происходят не в чане, где Йорик, а там, где Гамлет. Несмотря на то что мышление происходит «в чане», место личностных интересов – это местонахождение его тела.

Альтернатива 2. Где Гамлет, там и Деннет. Такое решение подходит сторонникам биологического подхода. И кажется, эта альтернатива ближе субъективному восприятию *Деннета*, созерцающего собственный мозг. Шекспировский герой, глядя на череп, обнаруживает бренность человека, а *Деннет*, глядя на собственный мозг... что это – не он! И правда, мозг, бездвижный, находится в чане, никуда, по-видимому, не собирается. А *Деннет* вот-вот отправится в Тулсу, где будет обезвреживать боеголовку. Или на секретное задание отправится не *Деннет*, а радиоуправляемый биологический робот? С одной стороны, интуиция, связываю-

⁴¹ О том, что сторонники тождества личности через тождество мозга являются фактически сторонниками психологического подхода, уже говорилось выше. См. также [Olson 1997].

щая личность с телом, с действующим в мире и воспринимающим агентом, слишком сильна. С другой стороны: разве нет противоречия в возможности воспринимать и быть воспринимаемым одновременно? Гамлетовские (йориковские?) размышления наводят героя на любопытную мысль, примиряющую, кажется, обе позиции.

Альтернатива 3. *Деннет* там, где он сам считает. Эта неожиданная новая позиция любопытна тем, что не предполагает жесткой зависимости между местоположением личности и какой-либо из ее физических частей. Точка зрения, а с ней и субъект, может мгновенно перемещаться с одного физического объекта на другой. В нашем случае – от мозга к телу. Более того, местоположение Я может вообще не соответствовать какой-либо физической части тела человека. Допустим, мы наблюдаем за событиями с помощью видеокамеры, находящейся в другом помещении. Если ничто не отвлекает нас, складывается устойчивое ощущение присутствия в этом другом помещении. Присутствия прямо за объективом видеокамеры. Значит, точка зрения может вообще не совпадать с положением частей человека. К этой позиции, как кажется, склоняется и герой мысленного эксперимента.

Но нас, читателей, она может насторожить. Совпадает ли на этот счет позиция *Деннета*-героя и *Деннета*-автора? В конечном итоге *Деннет* – один из самых последовательных сторонников материализма. Может ли он допустить, чтобы лич-

ность, подобно призраку, перескакивала с объекта на объект и даже существовала в пустоте? Дальнейшие зловключения героя мысленного эксперимента усиливают сомнения в полном совпадении взглядов героя и автора.

Итак, Йорик остается в безопасности, в стенах техасской лаборатории, а Гамлет отправляется к радиоактивному объекту. Но, не успев завершить задание, он попадает в беду: «Я совершенно оглох. Сначала я подумал, что сломались наушники моей рации, но когда я постучал по шлему, то ничего не услышал. По-видимому, сломались слуховые приемопередатчики... Я начал описывать то, что со мной происходило. В середине предложения я понял, что случилось еще что-то. Мой голосовой аппарат внезапно прекратил работать» [Ibid., 197]. Следом отказала правая рука, потом левая, и *Деннет* совсем ослеп. Проклиная ученых, впутавших его в эту авантюру, он сидит, оглохший, ослепший и неподвижный, в радиоактивной шахте, но вот... с отключением последнего передатчика он мигом переносится в Хьюстон. «[Е]сли минуту назад я был заживо погребен в Оклахоме, то теперь я был лишен тела в Хьюстоне» [Ibid., 197].

Моментальный перенос точки зрения становится поводом к «озарению» *Деннета*. Он столкнулся с подтверждением нематериальности души! «Ведь когда последняя радиосвязь между Тулсой и Хьюстоном прекратилась, я поменял местоположение со скоростью света! И моя масса при этом не увеличилась!.. Я не мог себе представить, как философ-физика-

лист мог бы это опровергнуть, не прибегнув к крайней, противоречащей здравому смыслу мере – запрещению всякого упоминания о личностях» [Ibid., 197–198]. Убеждение героя в субстанциональной нематериальной сущности Я вряд ли разделяет Деннет-автор. Но есть ли у последнего другой способ объяснить мгновенное перемещение личности героя? Пока мы размышляем над этим вопросом, герой «радовался своему философскому открытию и ломал голову... над тем, как сообщить о нем в журналы» [Ibid., 198]. К счастью, возможность установить связь с внешним миром и поделиться впечатлениями у него тоже вскоре появилась.

Конфигурация 2. Йорик (старый мозг) и Фортинбрас (новое тело)

Ученые погружают мозг *Деннета* в сон, а после пробуждения он обнаруживает, что обрел новое тело. «[П]осмотрев в зеркало, я удивился, увидав там незнакомца. Он был с бородой и немного толще меня» [Ibid., 198]. Да, Гамлет остался в подземной шахте, а у *Деннета* – новое тело. Его наш герой именует, конечно, Фортинбрасом. Оно также оснащено радиоконтактами. Поэтому Фортинбрас, как и Гамлет ранее, управляется Йориком дистанционно. Вскоре *Деннет* свыкается с новым обликом, изменившимся голосом, мускулатурой. Отражение в зеркале более не удивляет его – лицо несет все тот же оттенок светлого ума и решительного характера.

Ведь со сменой тела личность не меняется. И вот в один из дней он решает навестить Йорика, старого беднягу, который, как и прежде, находится в чане в Хьюстонской лаборатории.

«Я решил, что добрый старый Йорик заслуживает, чтобы его навестили. Я и мое новое тело... вошли в знакомую лабораторию под аплодисменты сотрудников» [Ibid., 199]. Перед *Деннетом* – прозрачный чан, в нем – мозг. А рядом – выключатель, который управляет системой радиопередатчиков между мозгом и телом. «[П]овинуясь случайному капризу, я протянул руку и перебросил тумблер в положение ВыКЛ. Представьте себе мое удивление, когда ничего особенного не случилось. Я не зашатался, не упал без сознания... – я не почувствовал ничего!» [Ibid., 199]. *Деннет* продолжает стоять напротив чана с собственным мозгом, смотреть на этот чан и размышлять, в то время как связь между ними, между Йориком и Фортинбрасом, нарушена. Как это возможно?

Конфигурация 3. Йорик (старый мозг) и Фортинбрас / Губерт (новый мозг) и Фортинбрас, Йорик = Губерт

Еще до того, как *Деннета* оперировали в первый раз, был изготовлен компьютерный дубликат его мозга. Была написана программа, которая работала в точности как независимая функциональная копия его мозга. Программу называли Губерт. Если на входе она получала аналогичные данные, сиг-

налы на выходе ничем не отличались от тех, что произвел бы мозг *Деннета* до первой операции. Программу подключили к радиопередатчикам Фортинбраса, так что теперь она получала всю ту же информацию, что и Йорик. Подобно мозгу, она обрабатывала данные и выдавала соответствующую реакцию. Только эта реакция не передавалась Фортинбрасу, а записывалась и сравнивалась. Наблюдение за синхронностью работы Йорика и Губерта длилось несколько месяцев, так что ученые смогли убедиться в абсолютной идентичности и независимости их работы. Тогда они решили совершить переключение. Когда *Деннет* прибыл в лабораторию, контроль над его телом уже осуществлял Губерт. Поэтому выключение радиопередатчиков Йорика никак не повлияло на поведение Фортинбраса.

В ходе мысленного эксперимента автор представил последовательно несколько возможных конфигураций субъекта: 1) *Деннет*, 2) Йорик-Гамлет, мозг и тело, соединенные дистанционно радиосигналами, 3) Йорик-Фортинбрас, мозг и новое тело, соединенные дистанционно с помощью радиосигналов. И вот теперь – 4) Йорик, Губерт и Фортинбрас: одно новое тело и два мозга – искусственный и органический, – работающие синхронно, одновременно получающие сигналы от органов чувств и попеременно выполняющие функцию управления организмом. Мыслительных центров у личности стало два, но, кажется, это не повлияло на количество личностей.

Изменения положения тумблера (Йорик→Губерт, Губерт→Йорик) проходят незаметно как для окружающих, так и для самого субъекта. Ведь точка зрения не смещается, и восприятия – одни и те же. «Я обнаружил, что, не считая щелчка, я не ощущал совершенно никакой разницы. Я мог повернуть тумблер в середине предложения, и фраза, которую я начал под контролем Йорика, завершалась без малейшей паузы под контролем Губерта» [Ibid., 200]. При изменении положения тумблера сохраняется тождество личности. Находясь и под управлением Йорика, и под управлением Губерта, *Деннет* не перестает быть одним *Деннетом*. У него остаются те же воспоминания, тот же характер, те же цели и желания. Значит, теперь у *Деннета* есть полноценный запасной мозг. *Деннет* – одна личность с двумя независимыми мозгами. «Я оказался обладателем искусственного мозга, который мог оказаться мне весьма полезен, если бы в будущем с Йориком что-нибудь случилось. Я также мог держать Йорика про запас и использовать Губерта. По-видимому, было совершенно безразлично, которого из двух я выбирал» [Ibid., 200]. Сюжетная линия мысленного эксперимента заставляет нас вообразить человека с двумя мыслительными центрами, работающими синхронно. И поверить, что они формируют одну личность.

Эта конфигурация напоминает трансплантацию психики. Но специфична тем, что вместо биологической модели реализации психика реализуется в программе компьютера. И

к тому же новый компьютерный мозг продолжает сохранять связь с Фортинбрасом. Это довольно экзотичный вариант. Но это – не последняя метаморфоза в злоключениях главного героя.

Конфигурация 4. Йорик и Фортинбрас / Губерт и Фортинбрас, Йорик ≠ Губерт; New fission case

Вполне ожидаемым продолжением сюжета могло быть наделение Губерта отдельным телом, скажем Розенкранца или Гильденстерна. Это было бы гуманно. Но автор избирает более экзотическую развязку. Это распад личности *Деннета* при сохранении одного тела. «СЛАВА БОГУ! Я ДУМАЛ, ЧТО ДО ЭТОГО ДЕЛО ТАК И НЕ ДОЙДЕТ! Ты не можешь себе представить, насколько ужасны были эти две последних недели – но теперь ты испытываешь это на собственной шкуре, поскольку наступила твоя очередь отправляться в чистилище» [Ibid., 201], – такие слова вдруг произнес *Деннет* после очередного переключения тумблера.

Дело в том, что около двух недель назад мозги перестали работать одинаково. Может, в код программы закралась какая-то ошибка или сбой произошел в радиопередатчиках, но состояния Губерта стали отличаться от состояний Йорика. «В мгновение ока иллюзия того, что я контролирую мое тело – наше тело[,] оказалась полностью нарушенной... Я чув-

ствовал себя так, словно меня носили в клетке, или, точнее, словно я был одержим. Я слышал, как мой собственный голос произносил слова, которых я не хотел говорить, я в отчаянии наблюдал, как мои руки делали то, чего я не хотел делать» [Ibid., 202]. Пятая, последняя представленная нам конфигурация: Йорик, Губерт и Фортинбрас. При этом Йорик и Губерт функционально различаются. В этой пятой конфигурации *Деннет*-герой, в представлении автора, распадается на две личности, два различных мыслительных потока, две независимые и несовпадающие воли, вынужденные делить одно тело. Хотя от органов чувств Фортинбраса на оба мыслительных центра все еще поступают идентичные сигналы, реакции Йорика и Губерта различны. И один из них вынужден быть узником, пока тумблер не перейдет в новое положение или пока ему не подберут новое тело. Рассказ заканчивается обещаниями одного *Деннета* найти подходящее тело другому. Ну а пока он (они) заняты этими поисками, мы попробуем добыть философские выводы.

История «Где Я?» – это художественное изложение проблемы тождества личности, или проблемы Исчезающей личности. И это мысленный эксперимент, в котором слышен голос самого агента, с которым происходят метаморфозы. История проблематизирует то, что кажется простым и очевидным, – местоположение личности. Деннету удастся представить это в виде проблемы потому, что он «разбирает» личность на модули и разносит их дистанционно. Но пробле-

ма, как было описано выше, в скрытом виде присутствует и в стандартной ситуации. Просто компактное расположение модулей прикрывает трудности локализации личности. Когда на скамейке сидят три человека, можно с большой вероятностью утверждать, что на скамейке находятся три личности. Но где эти личности находятся, с точностью сказать невозможно.

Стоит направить на человека увеличительные приборы, личность вдруг окажется невидимой и неуловимой, как и в деннетовской истории. Ведь даже в обычной жизни функции агента распределены, просто распределенность их не так масштабна. Дистанцирование мозга от тела заставляет героя и читателей задуматься: с каким из модулей правильной ассоциировать личность? Локализация когнитивных функций склоняет интуиции в пользу ассоциации с мозгом, а расположение органов чувств – в сторону ассоциации с телом. Но обе альтернативы, скорее всего, ложные. Это становится понятно, когда в конфигурацию распределенного агента добавляется запасной электронный мозг. Теперь альтернатив становится три, а в перспективе их бесконечное множество, ведь копий может быть сколько угодно. К тому же личность удивительным образом оказывается способна мгновенно перемещаться между всеми этими модулями. Как объяснить эти перемещения, не покидая пределов физикализма?

Еще более запутанной ситуация становится, когда мозги *Деннета* рассинхронизируются. Вдобавок к локализации те-

перь очевидной становится и проблема идентификации. Теперь не ясно не только где локализуется личность, но и кто из личностей – *Деннет*. Объяснить мгновенные перемещения героя можно, если считать личность не конкретной, физической, а абстрактной сущностью. Именно к этой идее подталкивает читателя автор. Абстрактные сущности нерезеренциальны, то есть не отсылают к каким-либо конкретным физическим объектам. И потому как их поведение в мире, так и их описание сильно отличаются от поведения и описания конкретных объектов. Меры длины, метры и километры, к примеру, необходимы для характеристики расстояния. Их использование вполне законно и не противоречит физикализму. Но было бы ошибочно искать эти абстрактные сущности где-то на дороге, считать количество таких характеристик или заключать об их тождестве между собой.

Атомы и пустота – не единственное, что признаёт наука. В ее онтологии есть место и для более странных вещей. Например, центр тяжести – понятие, которое составляет важную часть ньютоновской физики. Центр тяжести – не какая-то конкретная частица, не атом. Это условная точка, пространственная координата. У нее нет ни запаха, ни цвета, ни плотности, ни массы, нет никаких физических свойств. И тем не менее она необходима для того, чтобы точно предсказать движение объекта. Она даже фигурирует в каузальных объяснениях. В этом смысле она является реальным объектом, реальной объективной причиной событий. Центр тя-

жести может совпадать с геометрическим центром предмета, а может быть вне предмета. Центр тяжести Земли – где-то во внутреннем ядре, а центр тяжести бублика – в дырке. Перемещаться центр тяжести может не только постепенно, но и мгновенно, минуя промежуточные этапы. Такова экзотическая судьба абстрактных сущностей. Если следовать аналогии и предположить, что личность – это что-то вроде центра тяжести, можно без труда и без дуализма объяснить феномен, с которым столкнулся герой деннетовского мысленного эксперимента, – мгновенное перенесение личности. Впрочем, если личность сходна с абстрактным понятием, как могло возникнуть это абстрактное представление? В книге «Объясненное сознание» Деннет предлагает свою концептуальную схему возникновения личности.

2. Личность как результат эволюции

А. Появление границ, предпочтений и агентов

Некоторое время назад не было ни агентов, ни личностей. В природе были причины, но отсутствовали мотивы. Было разнообразие, но не было лучшего и худшего. Не было целей, так как некому их было иметь. Существенные изменения произошли с появлением химических структур, способ-

ных к самокопированию, – так начинает свою историю Деннет [Dennett 1991b, 173]. Это были простейшие живые организмы. Самокопирование было их ключевым свойством. Воспроизведение собственных копий стало их первым «интересом». Естественно, этот интерес у них был не осознан, но он объективно проявлялся в их поведении. Ведь выживать стали те, у кого способность к копированию была развита лучше. Они копировались чаще, и те копии, что производились с еще большей предрасположенностью к самокопированию, воспроизводились быстрее и в большем масштабе. То, что мешало воспроизводству и препятствовало работе механизмов самокопирования, стало «отрицательным», а то, что способствовало, – «положительным». Так появились предпочтения и «точка зрения». Конечно, эта точка зрения была внутренне пустой, «необитаемой». Но она уже существовала как некое место, с которого могла бы быть рассмотрена ситуация или... рассказана история⁴².

Примитивные системы, какими были первые репликаторы, не интериоризировали, не осознавали собственных предпочтений. Но то, что могло бы быть описано как их предпочтения, явно присутствовало в их поведении. Оно стало выражать рациональную схему. Поведение репликаторов возможно было интерпретировать как интенциональное, то есть

⁴² Подобную формулировку можно найти в «Размышлениях» Деннета к истории Р. Ракера «Программы» в сборнике «Глаз разума» [Хофштадтер, Деннетт 2003, 238].

преднамеренное. Рациональность появилась как объективная схема. Это была независимая от агента, свободная рациональность. И эта рациональность не могла существовать без определения границ между организмом и окружающей его средой.

Если главной задачей является производство как можно большего числа копий, сначала следует определиться, что нужно копировать, а что – нет. Невозможно же скопировать весь мир. К тому же для создания копий необходимо обеспечить стабильность, безопасность процессу копирования. Так возникла потребность в границах, разделяющих внутреннее и внешнее. «Как только твоим основным делом становится самосохранение, границы становятся необходимыми, поскольку, если ты хочешь сохранить себя, ты не должен растратить все силы, сохраняя весь мир» [Dennett 1991b, 174].

В мире неживой природы границ нет. Безразлично, где заканчивается один камень и начинается другой. Разлом или трещина могут проходить здесь или там – камни не извлекут из этого выгоду и не претерпят неудобств. Совсем другое дело – в случае живых существ. Они вынуждены следить за собственными границами. Иначе бы разрушилось их функциональное единство. Так что были созданы сложные механизмы защиты: шерсть, чешуя, панцири, зубы и клешни. Имунная система – это тоже инструмент защиты границ, целая армия из миллионов различных антител, чьей основной задачей является защита организма от незваных и вредных

для здоровья гостей. Правда, все механизмы защиты должны были быть направлены вовне, а не против себя. Как биологический организм мог определить, где его границы? Скорее всего, эта способность жестко запрограммирована эволюционным процессом. И она присутствовала у самых простых организмов. У всех организмов должно существовать рудиментарное представление о себе, о своей структуре. Иначе они не смогли бы адекватно реагировать на изменения во внешней среде. Эта рудиментарная схема и есть их собственное Я, простейшее биологическое Я.

В. Личность как расширенный фенотип

Существование границ и механизмов защиты у организмов не означало полной непроницаемости между внешним и внутренним. В организмах всегда были и есть пришельцы: это вирусы и бактерии. Эти пришельцы сами по себе являются автономными системами, у них есть собственные границы. Часть из них паразитирует на организме хозяина. А некоторые не только безвредны, но и необходимы для нормального функционирования последнего. Некоторые бактерии, например населяющие пищеварительную систему, позволяют усваивать пищу, а другие – вырабатывают полезные вещества. Даже митохондрии, основной источник энергии в наших клетках, скорее всего, некогда являлись «посторонними» – бактериями, которые были поглощены, но нашли

вариант взаимовыгодного сожительства с организмом-хозяином. Границы организма были проницаемы и в другую сторону – вовне. Ведь пространство агента может распространяться за пределы его физического тела [Ibid., 413].

Сами живые существа проводили экспансию во внешний мир. Многие виды приспособились осваивать окружающее пространство так, что оно как бы стало их естественным продолжением. Улитка накапливает кальций, чтобы потом выделять его вовне и формировать твердую ракушку, а рак-отшельник присваивает уже готовую раковину, пауки плетут паутины, а бобры строят плотины. Эти артефакты нужны им для того, чтобы повысить свои шансы на выживание, и, как показал биолог-эволюционист Ричард Докинз, они также являются объектами естественного отбора. Они представляют так называемый «расширенный фенотип».

«Расширенный фенотип» может продлевать территориальные границы организма на несколько сантиметров (как в случае с улиткой), а может – на километры (как в случае с границами заводи после строительства плотины бобром); он может быть относительно простым, а может быть необыкновенно искусно сделанным, как, например, термитник или гнезда австралийского шалашника. Он может включать в себя неорганические предметы, а может включать даже особей того же вида, как в случае с колонией муравьев. Но самые удивительные и сложные конструкции создаются человеком.

Мозг человека создает уникальную по своей сложности

паутину из поступков и текстов [Ibid., 416]. Как и искусственные артефакты животных, эта сеть – личная история – выполняет защитную функцию и увеличивает шансы организма на репродукцию. И как большинство артефактов животных, она собрана из элементов (и событий), уже существующих в окружающей среде. Это универсальная тактика самозащиты, самоконтроля и самоопределения. Язык и образованный на его основе биографический текст являются основой для появления личности – наиболее удивительного вида «расширенного фенотипа» у живых существ. Он создает личность как тренд, как геометрический центр биографических историй, как условную точку, куда указывает стрелка компаса поступков. Личность, согласно емкому определению Деннета, – это *центр нарративной гравитации*. «Наши истории выплываются, но в большей степени не мы плетем их, а они плетут нас. Наше человеческое сознание и наша нарративная личность – это их продукты, а не их источники» [Ibid., 418], – пишет американский философ.

Как и любое биологическое существо, человек и его мозг – это система из множества модулей. Ни один из этих модулей не обладает всеми уникальными качествами личности. Ни одному из них не могут быть приписаны желания, убеждения, действия агента и ответственность за них. Но в результате совместной работы всех этих модулей организм демонстрирует сложное целенаправленное поведение, которое и можно назвать поведением личности. Ее заметно с высо-

ты птичьего полета, с позиции автобиографической истории. Текст происходит как будто из одного источника, порождая ощущение единого агента, который создает эту историю.

История происхождения агентов, собственных Я, личности, в том виде, как эта история была представлена выше, конечно же, – набросок, модель, воспроизводящая логику эволюционного процесса. В этом сюжете нет подробностей и исторических дат. Тем не менее эта история позволяет наглядно проследить, как из неживой природы с помощью одного только слепого алгоритма могли возникнуть системы с собственной точкой зрения, границами, само-репрезентацией и даже автобиографией. Эта история показывает, что «личность» может быть продуктом биологического организма, результатом адаптации, лишь по уровню сложности, а не по принципу, отличающимся от других искусственных артефактов, созданных живыми существами. И возникновение такой личности естественно. Оно не требует осознанного плана по ее созданию, как плетение паутины не требует от паука интеллектуального замысла, как строительство термитника не требует архитектурного проекта у термитов, как работа иммунной системы не предполагает медицинских знаний у антител. Но из всего этого не следует, что личность – это что-то обыкновенное. В ее основе лежит рудиментарное представление о себе, саморепрезентация, которая присутствует почти у всех живых организмов. И она является одним из вариантов расширенного фенотипа, кото-

рый встречается даже у муравьев, бобров и пауков. Но при этом кажется, что личности все-таки принципиально отличаются от расширенных фенотипов других живых организмов. В чем заключаются эти отличия? В чем особенность личностей?

3. Критерии личности

«Вполне можно было бы надеяться, что столь важное понятие, применимость которого с такой уверенностью нами провозглашается или отвергается, имеет четко формулируемые необходимые и достаточные условия использования. Но если эти условия и существуют, мы их еще не выявили», — пишет Деннет в статье «Условия личностности» [Dennett 1976, 175]. Более того, философ считает, что их полный список в принципе не может быть составлен. Для того чтобы показать это, он идет «от противного». Деннет составляет список, который должен исчерпать критерии личности, а потом критикует его, показывая, что не только этот список, но и любой другой не сможет подойти для такой цели. Свой аргумент он начинает с общих наблюдений.

Класс личностей, замечает Деннет, по всей видимости, примерно совпадает с классом людей. Но связь этих классов не основана на определении. Личность — это понятие, относящееся не к биологическому виду, а скорее к функциональному. Другие биологические виды (и даже механические су-

щества) теоретически могли бы тоже обладать этой характеристикой. Однако требования к наличию личности, кажется, специфичны и высоки, так что другие существа, известные к настоящему времени, даже высшие млекопитающие, такие как собаки, обезьяны и дельфины, скорее всего, не соответствуют этим требованиям. Более того, даже некоторые люди не проходят по этим критериям – например, младенцы и люди, страдающие психическими расстройствами. Только к части людей мы относим понятие личности.

Кажется, критерии личности бывают двух типов: этические и метафизические. Уже философы Нового времени обнаружили это. Локк, в частности, писал, что личность – «это юридический термин, касающийся действий и их ценности и относящийся поэтому только к разумным существам, знающим, что такое закон, счастье и несчастье. Эта личность простирает *себя* за пределы настоящего существования, к прошлому только силой сознания; вследствие этого она беспокоится о прошлых действиях, становится ответственной за них, признавая за свои и приписывая их *себе* совершенно на том же самом основании и по той же причине, что и настоящие действия» [Локк 1985, 400]. Этический критерий личности – принадлежность к обществу морально ответственных агентов⁴³. Согласно этому критерию, если агент принад-

⁴³ То есть чтобы агент нес моральную ответственность, он должен быть личностью. Но «быть личностью» не означает быть морально ответственным за все поступки. Могут существовать условия, которые снимают с личности ответственность.

лежит к обществу морально ответственных агентов, то он – личность. С этим критерием нельзя не согласиться. Но он неинформативен. Получается круг: моральная ответственность предполагает существование личности, а личность понимается как то, что является основанием моральной ответственности. Для нетавтологического определения того, что такое личность, нужно использовать метафизические критерии.

Метафизические характеристики могут прояснить специфику личности. К метафизическим чертам личности Локк причисляет разумность, наличие сознания и памяти, знание добра и зла. Последнее можно называть моральной компетентностью. Во многом созвучные характеристики выделяет и Деннет в своей статье, посвященной понятию личности, «Условия личностности». Американский философ выделяет шесть основных условий: (1) рациональность, (2) наличие сознания, (3) возможность быть объектом личностной установки, (4) способность занимать личностную установку, (5) язык, (6) наличие самосознания [Dennett 1976, 177–178].

Разумность или рациональность у Деннета – это закономерность, телеологическая схема поведения, обусловленная эволюционным процессом. Она присуща множеству живых существ. Она предполагает наличие интересов и предпочтений у организма, а также минимальной способности эти потребности удовлетворять. Другая важная особенность – сознание. Это тоже общая черта для некоторых животных. Под

сознанием Деннет в данном случае понимает не качественные состояния, а только возможность приписывать организму интенциональные содержания: ментальные или психологические характеристики, желания и убеждения. Это функциональная характеристика, определяемая по поведению.

Следующее условие – способность быть личностью в глазах других. Деннет предлагает любопытную инверсию. Он считает, что не столько объективный факт того, что агент является личностью, заставляет других относиться к нему как к личности, а наоборот: относясь к кому-то как к личности, мы устанавливаем некоторое положение дел в отношении этого агента. «Считать ли нечто личностью, в некотором смысле зависит от *отношения, которое проявляют* к этому нечто, от *установки, которую принимают* по отношению к нему» [Ibid., 177].

Три первые характеристики, по мнению Деннета, выделяют класс, который он обозначает как класс интенциональных систем. «Интенциональные системы – это системы, чье поведение может быть (по крайней мере иногда) объяснено и предсказано на основании приписывания этой системе *убеждений и желаний...*» [Ibid., 179]. Класс интенциональных систем, по мнению Деннета, гораздо шире, чем класс личностей: в него попадают и животные, и даже некоторые машины. А вот следующие характеристики специфичны только для личностей.

Способность обходиться с другими как с личностями –

это еще одна важная характеристика. Это способность признавать наличие собственных интересов у другой личности, учитывать эти интересы при выборе своих действий. Чаще всего речь идет об уважительном отношении, поведении с учетом интересов и достоинства другой личности. Но не только. К личности также применимы и особые требования, в частности моральная ответственность. Даже некоторые формы преднамеренного крайне уничижительного отношения будут формой межличностных отношений. Личности по отношению к другим личностям могут проявлять особенную жестокость. Это тоже будет выражать специфически личностное отношение.

Одним из важнейших свойств личности является язык. Он дает возможность интериоризировать рациональность. В таком случае рациональность не только объективно прослеживается в поведении, в том, что существо имеет определенные цели и способно в поведении их добиваться, но и выражается в способности личность анализировать и планировать действия. Личности способны осуществлять предварительный отбор вариантов поведения внутри, в процессе рассуждений, и, по выражению К. Поппера, «позволяют гипотезам умирать вместо них самих» [цит. по: Dennett 1995, 375]. В отличие от других существ, выживающих только благодаря тому, что удачно выбрали вариант действия, личности не играют в рулетку – их выбор надежнее, чем надежда на случай. Если живые существа в целом просто удачливы благода-

ря имеющимся у них готовым адаптированным схемам, личности удачливы постольку, поскольку способны с помощью языковых выражений моделировать будущее.

И наконец, последнее, по мнению Деннета, условие – способность к самосознанию. Под самосознанием подразумевается представление о себе как об авторе собственных действий, способность описывать эти действия, наличие образа себя как длящейся во времени сущности. Эта способность напрямую связывает условие личности с условиями ответственности. В этом утверждении Деннет опирается на позицию Э. Энском, которая считала, что *«для того, чтобы быть ответственным за действие... необходимо осознать описываемое действие»* [Dennett 1976, 191]. Человек должен расценивать действие как свое собственное и иметь возможность объяснить его, исходя из собственной позиции. Ведь если он не может описать свое действие, опознать его в некоторой формулировке и объяснить его мотивы, значит, он не имеет особого привилегированного доступа к причинам действия, не может признать это действие своим и, соответственно, нет оснований считать именно его ответственным за это действие.

Приведенный выше список в целом характеризует способности обычного здорового взрослого человека. И представляется, что эти характеристики подходят на роль критериев личности. Все перечисленные условия необходимы для того, чтобы признать некоторый x личностью. Но, по мнению Ден-

нета, их все равно недостаточно, чтобы четко очертить класс личностей. Ни этот, ни какой-либо другой сходный список не позволяет этого добиться. Это связано с тем, что все перечисленные (и другие возможные) условия личности нормативны по своей сути. Они представляют абстрактный идеал, направление движения, но не позволяют провести четкую объективную разграничительную линию. «Люди или другие сущности могут только стремиться приближаться к идеалу, и невозможно установить какой-либо “зачетный балл”, который не был бы произвольным» [Dennett 1976, 193]. Рациональность, предсказуемость с позиции интенциональной установки, самосознание – все это идеальные модели. Ни одна личность не соответствует этим идеальным схемам все время, и поэтому наше суждение о принадлежности кого-то к классу личностей, к классу ответственных агентов все равно будет достаточно произвольным. «Личности в абсолютном смысле» не существует так же, как не существует «ответственности в конечном счете». Для того чтобы убедиться, насколько тонка грань между личностями и теми, кто находится вне этой категории, можно обратиться к случаям умеренных отклонений от норм: взрослых с признаками аутизма, умственной отсталости и психопатии.

А. Коллизия моральной и юридической ответственности в случаях психопатии

Особенно полезен в качестве иллюстрации феномен психопатии⁴⁴. Именно в отношении этого феномена проявляется коллизия юридической и моральной трактовок ответственности. Множество современных аналитических философов и психологов (Г. Ватсон, Д. Шумейкер, Н. Леви, С. Морс) считают, что психопаты не являются полностью ответственными моральными агентами. При этом в юридическом плане ответственность психопатов почти не подвергается сомнению. К психопатам применяются даже самые жесткие меры наказания, вплоть до смертной казни и пожизненного заключения. Значит ли это, что соответствующие законы сильно устарели и необходимо смягчить меры наказаний для психопатов? Или, наоборот, что критерии для «общества моральных агентов» философами применяются слишком строго? Наконец, может, допустима коллизия юридической и моральной трактовок ответственности?

Мне кажется, что этой коллизии быть не должно. Мо-

⁴⁴ Психопатия как термин имеет широкое значение. Так, в справочниках по психиатрии иногда насчитывается до 11 типов психопатии. Здесь в качестве примера я анализирую наиболее часто встречаемый в литературе (в частности, в работах наиболее известного канадского психолога Р. Хаэра) тип с характерными чертами. Этот тип психопатии лучше всего подходит для иллюстрации работы критериев личности Д. Деннета.

ральная и юридическая ответственность строятся на общих основаниях. В системе юридической ответственности наказание выполняет четыре основные функции: устрашение, ограничение дальнейшего возможного ущерба, реабилитация и воздаяние. Первые три функции имеют утилитаристское, прагматическое, направленное в будущее значение. Эти функции тривиальны и могут быть реализованы не только в отношении личностей, но любых агентов.

А вот последняя, четвертая функция наказания специфична. Кажется, чтобы ее реализовать, агент, в отношении которого она реализуется, должен обладать особенными свойствами. Он должен обладать компетенциями морального характера. Функция воздаяния имеет деонтологическое значение. Она направлена на установление справедливости как таковой, она направлена в прошлое. Но невозможно вершить справедливость в отношении того, кто абсолютно невосприимчив к категории справедливости. Аналогичное мы видим и в отношении моральной ответственности.

Моральная ответственность предполагает наличие у агента некоторой моральной компетентности. Соответственно, юридическая и моральная ответственность применимы в отношении одной особой группы агентов. Агенты, не подходящие для воздаяния, не могут быть ни морально, ни юридически ответственными. А значит, существующая коллизия должна быть разрешена. Но в чью пользу? Оправданно ли освободить психопатов от моральной и юридической ответ-

ственности, отказав им в статусе личности? Или нужно санкционировать применение этих видов ответственности в отношении психопатов?

В. Психопатия как расстройство личности

Психопатия – это разновидность «диссоциального расстройства личности». Это не столько болезнь, сколько тип характера. Диагностика психопатии осуществляется с помощью вопросника, который направлен на выявление определенных стереотипов поведения. Среди этих стереотипов – ненадежность, лживость, отсутствие переживаний вины и стыда, крайняя эгоцентричность и убежденность в собственной значимости, отсутствие склонности к глубоким и долгосрочным межличностным отношениям, отсутствие долгосрочного планирования, склонность манипулировать людьми, отсутствие сопереживания, импульсивность и проч. У психопатов отсутствует эмпатия – способность понимать чувства других и сопереживать другим людям. Психопаты, судя по всему, не переживают «сложных эмоций»: чувства вины, стыда, сожаления. Вследствие этого психопаты часто совершают аморальные поступки и преступления.

При этом основные когнитивные функции у этих преступников соответствуют норме. У них полностью развито мышление, память, существует система ценностей. Только система ценностей у них эгоцентрическая. Преступления психо-

паты совершают не под воздействием бредовых состояний, а полностью отдавая себе в них отчет, преследуя собственные цели. Мотивом антисоциального поведения психопатов является обогащение, повышение статуса или удовлетворение сексуальных желаний [Cornell, Warren, Hawk et al. 1996], то есть вполне «рациональные» соображения.

Причины психопатии изучаются. Этими причинами могут быть особенности воспитания и физиология. Среди нейрофизиологических причин наиболее вероятной является дисфункция миндалин и глазнично-лобных участков коры головного мозга [James, Blair 2003]. Эти участки отвечают за условно-рефлекторное отвращение и инструментальное обучение. И обе функции нужны для адаптации в обществе. Например, чтобы ребенок научился тому, что бить других – плохо, нужно, чтобы драка ассоциировалась у него с негативными реакциями: слезами, плачем, криками и страданием на лице жертвы. Так должно происходить отучение и от «аморальных действий» – через отрицательный стимул – отвращение к страданиям жертвы. Но у психопатов вид чужого страдания не вызывает отвращения. У них – дисфункция условно-рефлекторного отвращения. Поскольку они не испытывают отвращения к страданиям жертвы, они не могут научиться избегать подобных ситуаций. Они абсолютно глухи в отношении чужого горя и невосприимчивы к моральному закону. Поэтому с легкостью его переступают – конечно, когда считают, что это сойдет им с рук.

Отсутствие сопереживания. «Переживаю ли я за других людей? Это трудный вопрос. Ну да, конечно, я переживаю... но не даю чувствам быть помехой... Я, конечно, такой же сердечный и заботливый, как любой другой, но давайте говорить начистоту, все хотят тебя кинуть... Я должен позаботиться о себе сам, припарковать свои чувства, так сказать. К примеру, тебе что-то нужно или кто-то с тобой плохо обходится, хочет содрать с тебя двойную цену... приходится позаботиться о себе... сделать то, что должно быть сделано. Чувствую ли я себя плохо, если приходится причинять кому-то боль? Да, иногда. Но в большинстве случаев это как... эмм... [смешок] как ты чувствовал себя в последний раз, когда раздавил клопа» [Hare 1993, 33]. Это фрагмент из интервью, взятого у психопата, отбывающего срок за похищение людей, изнасилование и вымогательство, – назовем его Джо. В его собственных глазах он – самый обычный человек. И все его поступки – самые обычные и самые логичные. По его мнению, на его месте так поступал бы каждый. Просто ему не повезло в этот раз. Возможно, повезет в следующий.

У психопатов отсутствуют беспокойство, угрызения совести и стыд за причиненный ими вред. Их не волнуют последствия их действий для других. Участник ограбления, психопат, ранивший ножом случайную жертву, так комментирует свое преступление: «Да вы с Луны, что ли? Он проваляется несколько месяцев в больнице, а я гнию здесь. Я его немного порезал, но, если бы я хотел его прикончить, я бы перерезал

ему глотку. Я такой человек. Я его просто пожалел» [Ibid., 41]. Как видно, этот человек – назовем его Крисом – ни о чем не жалеет. Для него что сделано, то сделано, и на то были причины. И это типично для психопатов. А вот само откровенное признание – это скорее редкость. Об отсутствии чувства вины и сопереживания правильной всего заключать из повторяющихся паттернов поведения, бесконечных рецидивов преступлений. Ведь психопаты часто умело симулируют раскаяние. Ложность раскаяния обнаруживается только потом, в поступках и других высказываниях. Притворство, хамелеонство – одно из наиболее развитых качеств психопатов.

Хамелеонство. Р. Хаэр, американский психиатр, разработавший тот самый вопросник для диагностики психопатии, в своей книге «Лишенные совести. Пугающий мир психопатов» приводит [Ibid., 9–15] такой случай. В начале практики он устроился работать тюремным психологом. Его первым посетителем был Рэй. Рэй начал разговор с того, что попросил помощи: «Док, как дела? Слушай, у меня есть проблема. Мне нужна помощь. Я бы хотел поговорить на этот счет». После этого Рэй вытащил нож. Проблема Рэя была якобы в том, что к его «другу» приставал другой заключенный, и Рэй хотел тому отомстить. По этому поводу он, по его же словам, хотел посоветоваться с психологом. Но Рэй лукавил. В действительности он тестировал нового сотрудника Колумбийской тюрьмы. Заключенным запрещалось иметь оружие, и

перед Хаэром Рэй поставил непростой выбор: доложить начальству и разрушить свою репутацию перед всеми заключенными, или сохранить тайну и убедить Рэя не применять нож. Психолог выбрал последнее, и... нарушив тюремный режим, сам стал жертвой манипуляций Рэя.

Сначала тот убедил доверчивого психиатра, что его призвание – быть поваром. По ходатайству Хаэра Рэй был переведен для работы на кухню. Но поваром Рэй становиться не торопился. Получив доступ к запасам сахара, картошки и фруктов, он изготовил самогонный аппарат и стал производить алкоголь. Когда аппарат через некоторое время обнаружили, Рэй, как будто ничего не случилось, вновь появился в кабинете у Хаэра, на этот раз с просьбой о переводе в авторемонтный цех. Он, по его словам, чувствовал в себе способности мастера «и должен был готовиться к жизни за стенами тюрьмы». В конце концов Рэю опять удалось уговорить Хаэра. За месяц до ухода Хаэра с позиции тюремного психолога от Рэя поступила очередная просьба – рекомендовать его на работу к своему отцу. Рэй врал настолько искусно, что мог убедить любого, даже самого искушенного собеседника. К моменту знакомства с Хаэром за ним числилось множество преступлений, в том числе с применением насилия. Он половину жизни провел в тюремной камере, и тем не менее почти уговорил Хаэра. Когда тот все-таки отказал, Рэй разразился ругательствами. По дороге с работы у Хаэра в машине отказали тормоза – оказался надрезанным провод в тормоз-

ной системе. Накануне автомобиль ремонтировался в той самой мастерской, куда перевели работать Рэя.

Спустя несколько лет им предстояла еще одна встреча, когда Хаэр уже преподавал в Университете Британской Колумбии. В один из приемных дней он вдруг услышал уже почти забытый голос: «Да, я был помощником доктора Хаэра, когда он работал в тюрьме Британской Колумбии. Я помогал ему с документами, он обсуждал со мной все трудные случаи». Это был Рэй. Он пытался записаться на какой-то из курсов. Увидев Хаэра, он ничуть не смутился: «А, привет, док!» И тут же продолжил как ни в чем не бывало. Его не волновало, что ложь вскрывалась. Он просто менял тему или искал другую жертву.

Так поступают психопаты. Хамелеонство и лживость наряду с обаянием – их типичные черты. Они согласуются с их представлением о других людях как о средствах, инструментах в достижении корыстных целей. Возможно ли в таком случае относить психопатов к категории личностей? Как это соотносится с критериями, предложенными Деннетом?

По всей видимости, психопаты удовлетворяют большей части критериев, обозначенных американским философом. Они рациональны, сознательны, они свободно владеют языком (и владеют им иногда даже лучше нормальных людей), к ним относятся как к личностям. Вызывать сомнения могут только два пункта: (4) способность обходиться с другими людьми как с личностями, что является неотъемлемой

характеристикой личностей, и (б) самосознание. Сторонник лишения психопатов ответственности и статуса личностей мог бы сослаться на два этих пункта. Но мне кажется, что его позиция все же не имеет достаточно надежных оснований: психопаты в принципе понимают, как нужно обходиться с людьми, и осознают свои поступки в этическом плане.

В одном из примеров Джо явно показал, что достаточно остро переживает, когда с ним самим обходятся несправедливо, «хотят содрать с него двойную цену». И этим обосновывал свое жестокое ответное поведение. Судя по другим примерам, психопаты также способны понимать интересы и намерения других людей и даже иногда учитывают их в корыстных целях. Это видно, в частности, из ситуаций с Рэем. Так, он, манипулируя Хаэром, понимал, как, скорее всего, поведет себя новый тюремный психолог. Он догадывался, что Хаэр заинтересован в установлении доверительных отношений с заключенными даже больше, чем в соблюдении тюремных правил. И именно на этом строил свой план действий. Исходя из этого, можно заключить, что психопаты в целом способны понимать, как следует обращаться с другими личностями. Ведь для этого достаточно в принципе иметь способность понимать интересы других и знать, как к ним следует относиться. Способности же понимать, как следует обращаться с другими личностями, с моей точки зрения, достаточно, чтобы иметь способность относиться к другим как к личностям.

На это сторонник лишения психопатов статуса моральных агентов мог бы возразить: пусть так, пусть иногда психопаты способны относиться к другим людям как к личностям; но они совершают преступления, и именно в этих конкретных ситуациях они не способны видеть в других людях личностей. А значит, они не способны относиться к ним как к личностям и вести себя по отношению к ним иначе в момент преступления. Значит, в момент совершения преступления они сами не являются личностями, и это освобождает их от ответственности. Как личность может нести ответственность за поступки, которые были совершены не личностью, а более простым агентом?

Проблема с этим возражением в том, что «способность поступать иначе» в нем определена слишком узко. Сторонник аргумента пытается показать, что именно в данном конкретном случае в момент совершения преступления способности поступить иначе у преступника нет. Но мне кажется, что в лучшем случае он показывает, что в момент совершения преступления у преступника нет «возможности» поступать иначе, именно возможности, а не способности. По фактическим обстоятельствам мы не можем определять наличие или отсутствие способности. Способность определяется не фактическими обстоятельствами, а гипотетическими. К примеру, я сейчас сижу, но мог бы ходить. У меня есть способность как сидеть, так и ходить. Конечно, фиксируя все обстоятельства текущего момента, я не могу ходить, но на

основании этого нельзя делать вывод о том, что у меня в данный момент такой способности нет. Таким образом, я не вижу достаточных оснований считать, что психопаты не способны относиться к другим как к личностям. Тогда как обстоит дело с самосознанием?

Думаю, можно выделить две основные части этой компетенции: когнитивную и эмоциональную. Мне кажется, что приведенные выше примеры свидетельствуют о наличии у психопатов когнитивной, эпистемической моральной компетенции. Джо в принципе понимает, что значит быть сердечным и заботливым. Крис имеет представление, хотя и вполне ущербное, о жалости, а Рэй – о справедливости, воздаянии и соответствующем поведении. Психопаты знают о моральных нормах, ссылаются на эти нормы в речи и способны соотносить их со своими поступками и поступками других людей. В этом смысле они различают добро и зло, истину и ложь. Но соображения о добре и истине сами по себе не являются для них достаточным основанием поступать сообразно им. Поэтому скорее не когнитивный, а эмоциональный аспект моральной компетенции отсутствует или функционирует со сбоями у психопатов.

«Психопаты мало склонны испытывать эмоциональные реакции – страх и тревогу, – являющиеся ключевым механизмом в работе совести» [Ibid., 76]. У них как бы приглушен внутренний голос, останавливающий нормального человека от совершения антисоциальных поступков. Точнее,

этот внутренний голос лишен убедительности, которую придает ему эмоциональная окраска. К примеру, когда Крис намеревается ударить ножом свою жертву, он, вполне вероятно, осознает: «Если я его ударю, он будет страдать и истекать кровью». Но это осознание в эмоциональном плане для него почти равнозначно мысли: «Если я возьму такси, это будет несколько дороже, чем ехать на метро». Поэтому Крис не воздерживается от преступления.

Убеждение, что у психопатов нарушен именно эмоциональный, мотивационный аспект осознания, разделяют много исследователей. Среди них – Д. Шумейкер [Shoemaker 2007], Дж. Мёрфи [Murphy 1972] и Э. Дафф [Duff 1977]. В силу этого порока они считают психопатов неподходящей мишенью для моральных оценок и несоответствующим объектом для межличностных отношений. Шумейкер так аргументирует свою позицию. Для того чтобы обладать моральной компетенцией, быть объектом для действия моральных законов, агент должен быть способным действовать исходя из моральных соображений. Долженствование предполагает возможность. Психопаты способны совершать поступки, которые только выглядят как моральные, но их мотивация в любом случае эгоистична. Действия из моральных соображений предполагают способность к эмпатии, а она отсутствует у психопатов. Психопат не может искренне заботиться о другом человеке. Его интерес к другим исключительно «инструментальный». Ценность других людей для него

определяется их полезностью. Поэтому психопаты неспособны действовать из моральных соображений. Они не обладают моральной компетенцией. В этом смысле они не участвуют в межличностных отношениях и, возможно, отстают даже дальше от нормальных людей, чем некоторые животные. Шумейкер иллюстрирует этот аргумент воображаемой ситуацией.

С. Моральное чувство инопланетян

Предположим, среди людей поселяются инопланетяне. Их моральные принципы в целом соответствуют человеческим, но распространяются на более широкий класс объектов. В частности, инопланетяне считают аморальным ходить по траве, так как трава от этого получает ущерб. Нанесение вреда траве само по себе плохо, считают инопланетяне и мотивируют это тем, что знают, «что значит быть травой», которой нанесен физический ущерб. Когда люди им возражают, обращая внимание на то, что трава не обладает сознанием, инопланетяне отвечают, что это не важно. Для того чтобы знать, что значит быть травой, нет необходимости переносить на себя ее ментальные состояния. Достаточно «войти в ее пространственное существование». Для людей эти основания, безусловно, представляют бессмыслицу.

Предположим теперь, что какой-то человек прогуливается по парку и видит вдали любопытную формацию горных

пород. Из любопытства он решает подойти к ней поближе, чтобы лучше ее разглядеть. Приближаясь к формации, человек наступает на траву. Он не может испытывать эмпатии к траве, поэтому не чувствует аморальности своего поступка. Но тут к нему в ярости приближается инопланетянин: «Как плохо вредить траве!!! Как ты мог совершить такой аморальный поступок?!!» Есть ли основания у инопланетянина считать человека морально ответственным за нанесение вреда траве? Нет, считает Шумейкер, ведь люди просто не способны понять, какой ценностью обладает неприкосновенность травы. Сколько бы инопланетяне ни растолковывали людям о «вхождении в пространственное существование травы», это никак не убедит людей в их «моральном долге» перед травой. И психопаты – в таком же положении. Они неспособны воспринять мотивирующую силу моральных норм в отношении других. Кажется, что психопат в обществе других личностей должен чувствовать себя, как человек среди инопланетян со сверхчувственной моральной компетенцией: он должен понимать требования окружающих, но не видеть у них оснований. И при первом возможном случае нарушить правила, если это останется безнаказанным. Психопат не обладает моральной компетенцией в отношении других людей и потому не является полноценной морально ответственной личностью.

Я не согласен с такими выводами из вышеописанного случая. Интуиции об отсутствии моральной ответственности в

данной истории основаны на иррациональности требований инопланетян Шумейкера. «Пространственное существование» растений не удовлетворяет критериям рациональности оснований морального долга. Если под аналогичные требования подвести рациональные основания, даже если субъект действий не будет обладать соответствующим чувством эмпатии, мысленный эксперимент вызовет интуиции о наличии у него моральной ответственности. Для того чтобы это подтвердить, я предлагаю сравнить историю Шумейкера с новой историей. В ней требования рациональны, но субъект не испытывает эмпатии. Это – моя история про общину джайнов.

Д. Община джайнов и ахимса

Молодой человек присоединяется к общине джайнов. Философия и правила жизни джайнов направлены на совершенствование души и освобождение от круговорота рождений и смертей. Путь к совершенству лежит в соблюдении пяти принципов: быть искренним и благочестивым, не красть, не прелюбодействовать, не стяжать и не причинять вреда живому (ахимса). Из них пятый принцип – самый важный. Он запрещает джайнам наносить любой вред любому живому организму, даже насекомым и растениям. В основе этого принципа – убеждение в том, что любое существо наделено индивидуальной и вечной душой (дживой), способной к

страданию и обладающей абсолютной ценностью. Соблюдая этот принцип, джайны трепетно относятся ко всему окружающему. Они избегают передвигаться ночью или даже пить непроцеженную воду, так как боятся случайно нанести вред крохотным организмам.

Молодой человек хорошо знаком с порядками в общине джайнов. Но, возвращаясь в одиночестве домой, он решает пренебречь обетами общины. Вместо того чтобы следовать по окружной дороге, он решает пойти напрямик и, не глядя под ноги, пересекает луг, протаптывая себе тропинку. Он не чувствует мотивации соблюдать обеты джайнов, не испытывает эмпатии к насекомым и растениям, которые могут оказаться под его ногами. Поэтому идет без всякой осторожности. В этот момент его окликает учитель. Он укоряет юношу за нарушение заповедей. Он считает, что юноша виноват в несоблюдении правил общины. Оправдан ли в данном случае укор?

Мне кажется, укор и осуждение вполне здесь уместны. Уместно говорить о моральной ответственности юноши. Она уместна, даже если молодой человек не испытывает мотивации и не чувствует эмпатии ко всем живым существам. Присоединяясь к общине джайнов, он принимает их установки в качестве минимальных моральных требований. К тому же он рационально вполне может понять природу их требований. И то, что молодого человека больше мотивирует желание сэкономить время, чем обеты джайнов, не снимает с него от-

ветственности. Этот пример является доводом в пользу моральной ответственности психопатов. Психопаты среди нормальных людей – как молодой человек среди джайнов, а не как люди среди инопланетян Шумейкера.

Впрочем, чтобы убедиться в наличии моральной ответственности психопатов, мне кажется необходимым привести еще один аргумент. Этот аргумент построен на аналогии между психопатами и закоренелыми преступниками.

Е. Психопаты и закоренелые преступники

Сторонник освобождения психопатов от моральной ответственности считает, что психопаты лишены способности обходиться с другими как с личностями и/или не обладают моральной компетенцией, так как постоянно совершают жестокие преступления. Именно из-за регулярности совершаемых преступлений они не могут быть морально ответственными. Но этот аргумент верен и в отношении любых закоренелых преступников. Ведь они тоже постоянно совершают жестокие преступления. Значит ли это, что они тоже должны быть освобождены от моральной, а также и уголовной ответственности? Если моральная компетенция снижается пропорционально частоте и тяжести аморальных поступков, самые закоренелые и жестокие преступники оказывались бы за рамками морального осуждения. Но тогда сама практика моральной критики и наказаний потеряла бы всякий смысл.

Это – сведение к абсурду первого аргумента. Очевидно, закоренелые преступники должны подвергаться осуждению и должны быть уголовно наказуемы. Следовательно, и психопаты должны подвергаться такому осуждению. И дело здесь даже не столько в наличии способностей или их отсутствии, сколько в чертах характера и качестве воли преступников. И психопаты, и «психически нормальные» закоренелые преступники склонны к регулярному совершению преступлений. И те и другие обладают скверным характером и являются аморальными личностями. И те и другие обладают «злой волей», то есть намеренно причиняют зло другим. И способность к эмпатии или ее отсутствие здесь не так важны.

Что дает наличие самой возможности эмпатии? Сравним аморальные действия двух человек. Предположим, они совершают одинаковые действия с одинаковыми мотивами – мошенничеством похищают деньги у пожилой женщины. Один человек, более жестокий, никогда не испытывает эмпатии, а другой – переживает ее в редких случаях. Является ли особенность характера первого основанием освобождения его от ответственности? Мне кажется, что нет. Во-первых, у преступников – одинаковые мотивы. Эти мотивы основаны на личной выгоде. И действия совершены не неосознанно, а именно на основе этих мотивов. Во-вторых, оба преступника продемонстрировали «злую волю» – в поведении их проявилось суждение «мои интересы – превыше всего». Тогда почему только второй должен нести ответствен-

ность? Почему нереализовавшаяся альтернативная возможность должна повлиять на оценку актуализованного осознанного действия?

На мой взгляд, если и есть разница в оценке поступков двух этих людей, она – не в пользу первого преступника. Первый является жестокосердным по натуре. И это – лишь дополнительное основание для морального осуждения. Представляется, что основаниями ответственности служат не столько альтернативные возможности, сколько актуальные мотивы и ценности. Следовательно, и психопаты должны нести моральную ответственность. Возможно, даже в большей степени, чем преступники с нормальной психикой. Психопаты и закоренелые преступники различают добро и зло как минимум в том смысле, что могут определять, что интерпретируется обществом как зло, а что – как добро, они способны сопоставлять эти нормы со своими планируемыми действиями и вступают в обществе в межличностные связи, поэтому они являются личностями и несут моральную ответственность.

Впрочем, этот спор показывает нам, что границы класса личностей размыты. Можно оспорить соответствие психопатов двум необходимым требованиям, предъявляемым к личности, как это было рассмотрено выше. Но можно также оспорить соответствие каждого человека, совершающего антисоциальные поступки, этим и другим требованиям. И тогда, если признать этот ход корректным, все злоумышленни-

ки окажутся освобожденными от ответственности. Это явно абсурдная ситуация. Следовательно, к каждому критерию нужно применять оператор «отчасти»: личность – это человек «отчасти разумный», «отчасти сознательный», «отчасти способный относиться к другим как к личностям», «отчасти обладающий самосознанием» и «моральной компетенцией». Без этого оператора категория личностей останется пустой. Как и моральная компетенция, остальные критерии являются идеальными конструкциями, которым нельзя подобрать точные соответствия в мире. Рациональный, самосознательный, морально компетентный и даже просто сознательный – это абстрактные характеристики, ориентиры, а не реальные свойства людей. Агент в различное время оказывается дальше или ближе в отношении этих ориентиров, но никогда не достигает их в полном объеме, тем более во всех своих действиях. Таким образом, личности «в конечном счете» не существует и существовать не может, кроме как в метафизических фантазиях. Но это не лишает смысла рассуждения о личностях.

Пусть понятие личности – нормативная модель и представляет собой слишком высокую планку для биологических организмов. Но это не отменяет практическую пользу этого понятия. Если бы нормативные понятия вообще не имели смысла, не имели бы смысла и все этические построения. Зато теперь мы знаем, какого рода сущность следует искать. Это знание поможет нам в решении вопроса о реидентифи-

кации и атрибуции ответственности. Это поможет, в частности, принять некоторые контринтуитивные положения нарративного подхода – наиболее перспективной теории личности.

V. Нарративный подход к вопросу тождества личности

Нарративный подход отличается от подходов, которые мы исследовали раньше. Первое важное отличие – в том, что большинство его сторонников не считают характеристики и свойства личностей объективными. Согласно мнению нарративистов, вопрос о характеристиках и свойствах личности, а также о ее тождестве во времени всегда зависит от наблюдателей и рассказчиков. «Нарративный подход отказывается от попыток отождествления “я” с некоторым ментальным или физическим объектом в мире. То, на что указывает местоимение “я”, – это конструкция, возникающая в подходящем социально-биологическом контексте» [Левин 2018, 185], – пишет российский исследователь и аналитический философ С. Левин. То есть первичным по отношению к личности является рассказ, а не наоборот. Рассказ порождает личность, а не личность – рассказ. Кроме того, этот рассказ постоянно меняется и дополняется наблюдателями, соответственно, параметры личности никогда не являются установленными до конца. Из всего этого следует, что нарративный подход – это скорее антиреалистическая концепция личности. Этим он отличается от реалистических психологического, субстанционального и биологического подходов.

Второе важное отличие – в самой постановке вопроса. Как

я уже писал, сторонники психологического и субстанционального подходов заняты вопросом *реидентификации* личности в различные моменты времени. Они пытаются определить, что делает одну личность в момент t_2 той же самой личностью, что существовала в момент t_1 (или, точнее, определяют, что делает одну временную стадию личности тождественной другой временной стадии личности). Сторонники биологического подхода *реидентифицируют* личности с объектами или животными. То есть философы классических направлений устанавливают *отношения нумерической тождественности* между временными стадиями личностей (*person time-slices* или *person-stages*) или между личностями и объектами. Их оппоненты, сторонники неклассического, нарративного подхода, в частности М. Шехтман, Д. Деннет и К. Эткинз⁴⁵, не считают отношение нумерического тождества существенным для практических целей.

Для объяснения сохранения длительного существования личности и обеспечения условий моральной ответственности (как, впрочем, и для критериев выживания и справедливой компенсации), по их мнению, нужно решать не вопрос реидентификации, а вопрос *характеризации личности*⁴⁶. Для того чтобы определить степень ответственности

⁴⁵ Другими сторонниками нарративного подхода являются, например, Тейлор, Веллеман и Макинтайр.

⁴⁶ Такая постановка вопроса особенно характерна для М. Шехтман [Schechtman 2007].

личности, не нужно искать нумерического тождества или сравнивать стадии личности, а нужно выявить, какие психологические характеристики, ментальные события и физические действия могут быть ей приписаны. А для этого следует исследовать моральную сущность личности, определить, «какова эта существующая личность», какие черты для нее являются ключевыми, а какие – случайными. Таким образом, сторонники нарративного подхода заняты отношением принадлежности личностям соответствующих свойств, состояний и действий, а не отношением нумерического тождества.

Принадлежность психологических характеристик, ментальных событий и физических действий личности в нарративном подходе определяется возможностью гармоничного и непротиворечивого включения их в биографическую историю личности как принадлежащих самой личности. Под биографической историей понимается обычный последовательный рассказ о развитии личности, преимущественно от первого лица. Именно возможность непротиворечивого включения в последовательную реалистичную автобиографию делает эти события и характеристики частью психологической сущности и подлинной идентичности личности, а также является основанием ее моральной ответственности за соответствующие поступки. Нарративный подход кажется наиболее близким обыденной практике вынесения моральных суждений.

Допустим, вы пытаетесь разобраться, несет ли ответственность человек за какое-то происшествие. Конечно, можно опираться на одни только улики, вещественные доказательства и факт присутствия человека на месте происшествия. Но логичней всего задать ему и свидетелям событий вопрос об этом происшествии. В ответ, скорее всего, последует разъяснение в виде истории. Агент расскажет о том, что произошло, о чем он думал, на что надеялся, что получилось. По рассказу можно будет понять, какое отношение имеет этот человек к происшествию, какие мотивы им двигали, какую выгоду он приобрел и какие потери понес. Только по истории, в которой объединятся перспективы разных участников и свидетелей, можно будет определить вину или невиновность того или иного лица. Без истории нельзя определить, случилось ли убийство или это была самозащита, что было случайным стечением обстоятельств, а что – реализацией хорошо продуманного плана. Без истории нельзя ни разобраться, что произошло, ни справедливо оценить происшествие.

1. Конструкция характеризующего биографического нарратива

Но как определить, что может, а что не может быть непротиворечиво включено в историю, в биографический контур нарратива? Что может объединять историю и служить фильтром для нерелевантных элементов? Во-первых, историю

может объединять *перспектива*. Истории обычно рассказываются или от лица какого-то персонажа, или поочередно от лиц нескольких героев, или с позиции автора. Но в каждый момент эта перспектива есть. Перспектива – это взгляд с точки с четкими временными и географическими координатами. Перемещение этих точек в развивающейся истории в основном последовательно и непрерывно как в пространстве, так и во времени. Даже тогда, когда это перемещение скачкообразно (как в истории «Где Я?»), скачки имеют объяснения.

Конкретное расположение точки перспективы обуславливает избирательное отношение к обстоятельствам и выражается в описаниях под каким-то специфическим углом зрения. Даже когда это перспектива с высоты «птичьего полета», это – конкретная перспектива. Она действует как прожектор, выхватывая из бесконечного множества допустимых описаний основное. Исходя из некоторой перспективы, часть обстоятельств выделяется как значимая, важная, основная, а другая часть, собственно большая часть, попросту игнорируется. Значимые события изображаются подробней, их временные рамки как бы раздвигаются. А незначительные обстоятельства упоминаются мельком или вообще не оговариваются. Отношение между различными по важности планами систематично. Выделяются главные и побочные персонажи, ключевые, образующие историю события и случайные обстоятельства. Ключевые события образуют канву сюжета

или вехи биографии. Таким образом, законы перспективы позволяют отобразить и правильно расположить элементы рассказа.

Второе возможное объединяющее начало – это *траектория психической жизни персонажа*. История обычно состоит из объективных событий и действий героев. Действия, как мы уже знаем, увязываются с ментальными состояниями, а эти ментальные состояния, сочетаясь между собой, согласуются с сущностными характеристиками личностей. Ментальные состояния объясняют поступки, а характер и склонности личности – ментальные состояния. Из одних ментальных состояний проистекают другие. Таким образом, вся жизненная траектория должна быть в истории интеллигибельна. Это не означает, конечно, неизменности характера. Обстоятельства истории влияют на героя, они могут значительно изменить его состояния, характеристики и даже сущность. Но все эти изменения происходят не без причин. Причины включаются в историю и объясняют эти изменения. Таким образом, вся история становится понятной, последовательной, осмысленной, а траектория психической жизни персонажа не перестает быть непрерывной. В этой интеллигибельности – второе основание единства и фильтр для элементов нарратива.

Третий скрепляющий элемент биографической истории и одновременно ее фильтр – *телеологическая направленность*. История может объединяться движением личности к

каким-то долгосрочным целям. Иерархия целей может увенчиваться каким-то основным ориентиром, главным маяком в путешествии личности. Именно этот главный ориентир, конечный смысл в масштабе всей жизни наделяет значением отдельные поступки и состояния. Кажется, главные ориентиры могут быть разные. Это может быть выживание, личное благополучие, доминирование, навязывание другим своей воли или получение максимального удовольствия. Но некоторые сторонники нарративного подхода (в первую очередь герменевтического направления) настаивают, что подлинное «личное благо» должно совпадать с добром как всеобщей ценностью: желанием жить хорошо вместе с другими и для других в справедливом обществе. Только наличие этого главного этического ориентира, по их мнению, делает осмысленной жизнь в целом и каждое действие в отдельности.

Перечисленные выше характеристики позволяют изнутри скрепить нарратив, сложить жизненную историю в одно целое. По отношению к содержанию они являются *внутренними* требованиями соответствия логике и структуре биографической истории. Но есть еще два необходимых *внешних* условия, предъявляемые к биографическим нарративам для атрибуции характеристик агентам. Это условия *возможности изложения и правдоподобия нарратива*. Биография может быть последовательна в отношении перспективы, представлять ясный психологический портрет героя, быть осмыс-

ленной, обладать телеологичностью, но при этом не давать достаточных оснований для атрибуции действий и моральной ответственности за них. Это – случаи, когда сам агент не может сформулировать и выразить собственный биографический нарратив или его часть или когда нарратив целиком представляет собой фикцию. В первом случае агент может быть неспособен в достаточной мере анализировать и оценивать свое поведение, а во втором его оценки могут быть далеки от реальности.

Способность сформулировать *полную* жизненную историю в рассказе от первого лица, удовлетворяющем внутренним требованиям, не является абсолютной необходимостью для атрибуции действий и ответственности. В конце концов, агент может не понимать мотивов своего поведения, не всегда различать конечные цели, даже не иметь возможности последовательно и подробно изложить события всей своей жизни. Но важно, чтобы агент имел способность организовывать свой опыт и представления в жизненную историю. Свидетельством ее наличия является способность агента выражать в истории хотя бы часть жизненных событий и объяснять их. Я имею в виду способность описывать свои поступки от первого лица, ссылаться на их ментальные причины и обосновывать свои суждения об опыте. Если агент способен к такому описанию фрагментов жизни, значит, он способен структурировать свой опыт в виде нарратива, и, следовательно, то, что вписывается в этот нарратив и отвеча-

ет минимальным требованиям реализма, может служить его характеристикой.

Правдоподобие биографической истории – второй, но не менее важный внешний критерий. Требование правдоподобности, как и требование наличия способности описать опыт, также не может быть абсолютным. Повествующий агент может допускать и даже всегда допускает какие-то ошибки. Но хотя бы часть суждений о событиях в нарративе должна совпадать с суждениями других личностей. Иными словами, описания событий от первого лица должны хотя бы частично совпадать с описаниями от третьего лица других участников и свидетелей событий. Иначе это будет фантастический нарратив, и его нельзя будет использовать в качестве основания атрибуции действий и ответственности. Некоторые причины тому обсуждались еще во Введении. Выше в тексте этой главы также было показано, что людей, представления которых полностью оторваны от реальности, вообще невозможно относить к категории полноценных личностей. Даже если они способны понимать этические модели поведения (в том числе способны представлять, как можно относиться к другим как к личностям), они неспособны применять эти модели на практике. А без этого они не будут удовлетворять требованию моральной компетенции. И фантастическая автобиография такого человека не будет удовлетворять требованиям нарративного подхода. Работу внутренних и внешних критериев можно продемонстрировать на примере сле-

дующей ситуации.

Предположим, кто-то ошибочно называет себя Наполеоном и утверждает, что был командующим в знаменитой битве у Ватерлоо. Участие в этой битве он обосновывает стремлением восстановить величие Французской империи, а за разгром считает ответственным лично себя. Теперь его целью является мировое признание и получение должного императору уважения. Как соотносится такая история с требованиями к нарративу? История артикулирована. В ней есть некоторая телеологическая целостность – стремление к мировому признанию. Но остальным критериям она не удовлетворяет. При анализе этого нарратива возникнут несоответствия. Нарушена будет непрерывность точки перспективы. Например, заблуждающемуся человеку сложно будет объяснить, что он делал в период между 1821 и 2018 гг. и как сочетаются его биография и ментальная жизнь с фактами жизни Наполеона. Скорее всего, нарушится требование к интеллигентности. Скажем, этому человеку будет трудно объяснить, зачем в XX в. он поступал в начальную школу, если в конце XVIII в. уже заканчивал кадетский коллеж. Но основной будет проблема несоответствия реальности. Так, этому человеку будет трудно объяснить физические различия между Наполеоном и им самим. Таким образом, его история не будет соответствовать критериям интеллигентности, единства перспективы и правдоподобия. Следовательно, она не сможет быть характеристикой этого человека и осно-

ванием для атрибуции ему описываемых действий. Совсем другая характеристика будет более подходящей для этого человека: что он – вдохновленный историей фантазер или психически больной.

2. Преимущества нарративного подхода

Нарративный подход предполагает, что личность формируется из рассказа и, соответственно, обладает характеристиками, расположенными в разных временных координатах. Личность реализуется в жизненной истории, автобиографии. В этом и есть сущность личности. Поэтому для личности особенно важно, чтобы нарратив развивался максимально благополучно. Личность имеет основания ожидать справедливой компенсации в рамках этого нарратива и несет ответственность за действия, включенные в нарратив. Особую роль в формировании этого нарратива играет перспектива первого лица. То, что рассказ из этой перспективы имеет большое значение, дает дополнительные автономии, контроль и свободу личности⁴⁷. При этом автономия личности не абсолютна. Автономия ограничивается внутренними и внешними критериями характеризующего автобиографического нарратива. Имеет ли этот подход преимущества в срав-

⁴⁷ Многие философы думали, что человеческая автономия включает или, возможно, даже состоит из способности самосозидания – способности определять, изобретать и создавать себя [Velleman 1996].

нении с другими подходами? И решает ли он проблему, обозначенную в Аргументе пяти трансплантаций?

Аргумент пяти трансплантаций показывает, что субстанциональный, биологический и психологический подходы не объясняют *особой* заботы личности о будущем. Забота о «своем» будущем, в соответствии с этими подходами, в целом не должна отличаться от заботы о будущем других личностей. Парфит называет это утверждение *Крайним утверждением (Extreme claim)* [Parfit 1986, 307]. Крайнее утверждение следует из Аргумента пяти трансплантаций. Обратной стороной Крайнего утверждения является то, что я назову *Исключающим утверждением*. Исключающее утверждение состоит в отрицании возможности атрибуции моральной ответственности ввиду отсутствия асинхронного тождества личности, исчезновения агента. Нарративный подход, определяя личность не как последовательность временных фаз, а как распределенный во времени рассказ, решает эту проблему и, следовательно, обходит как Крайнее, так и Исключающее утверждения. Этот подход опирается на идею фундаментального единства нарратива и его частей. И благодаря этому способен объяснить особую заботу личности о своем будущем и ее ответственность за свое прошлое.

Субстанциональный, биологический и психологический подходы оказываются уязвимыми для гипотетических случаев с делением и умножением личности. Следствием применения этих подходов к подобным ситуациям является про-

тиворечие или эпистемологический пат. У сторонников этих подходов есть только две альтернативы. Либо им необходимо признать, что определить наследника существовавшей до деления или умножения личности невозможно. Либо что одна личность вдруг становится тождественной нескольким разным личностям. Нарративный подход обходит эту дилемму. Это оказывается возможным потому, что связь через нарратив фиксирует отношение принадлежности, а не отношение асинхронного тождества. Предположим, произошло деление или умножение личности. Согласно нарративному подходу, наследниками действий и моральной ответственности одной личности будут несколько человек. У этих людей будет одинаковое прошлое. Но в данном случае противоречия не возникнет. Как стена может быть частью и гостиной, и спальни одновременно, так и часть нарратива может входить в разные биографии. Следовательно, гипотетические случаи с делением и умножением личностей не являются проблематичными для нарративного подхода. И этот подход обладает преимуществами по отношению к другим подходам. Но насколько он уязвим для критики? Кажется, некоторые варианты возражений, уже опробованные выше, могут быть применены и к нарративному подходу, в частности возражение, основанное на разнообразии модусов человеческой жизни.

3. Возражения к нарративному подходу

Возражение 1. Проблема двойников

Одно из самых существенных возражений к психологическому критерию тождества личности строилось на возможности возникновения двойников, тождественных в качественном смысле, но различающихся нумерически. Эта возможность была продемонстрирована в Аргументе редупликации, а также с помощью примера Телетранспортер. Однако может показаться, что аналогичное возражение работает и в отношении нарративного подхода. Такое возражение высказывает, например, российская исследовательница М. Секацкая [Секацкая 2018]. Она считает, что у нарративного подхода даже меньше шансов справиться с этой критикой, чем у психологического. Как объяснить возможность существования нескольких личностей с одинаковыми нарративами в случае с редупликацией? Кому приписывать ответственность и как решать вопрос о выживании в случае с ветвлением личности?

Я считаю, что это возражение важно, но не из-за того, что представляет угрозу для нарративного подхода, а из-за того, что позволяет его лучше прояснить. Ситуации ветвления личности допустимы для нарративистов и не пред-

ставляют проблемы именно благодаря тому, что сторонники нарративного подхода для решения практических вопросов не считают необходимым соблюдение нумерического тождества. Они не занимаются нумерической реидентификацией личностей в разное время, а ищут способы характеристики личности. Следствием такого подхода является возможность пересечения характеризующих нарративов, ветвление и объединение их. Возьмем в качестве иллюстрации уже известный нам случай с Гаем Фоксом.

В описанном выше сценарии дубликации Роберт и Чарльз становятся Гаем Фоксом, обретая его воспоминания, характер, убеждения и привычки. Они оба утверждают, что они Гай Фокс. Более того, сторонние наблюдатели не могут отрицать, что по всем данным Роберт и Чарльз являются Гаем Фоксом. В данном случае нарративный подход не отрицает этого факта. Согласно этой концепции, оба человека несут ответственность за действия Гая Фокса и оба являются продолжениями его личности. Пока их история совсем неразличима – это одна личность с двумя телами. Однако это состояние краткосрочно, так как истории неизбежно разойдутся, и тогда личностей будет две. Дубликация личностей не ведет к проблеме, ведь характеристика, в отличие от нумерического тождества, не является отношением один к одному. У двух разных объектов могут быть общие характеристики так же, как у двух разных объектов могут быть общие части. Таким образом следует объяснять дубликацию в нарратив-

Возражение 2. Личность предшествует нарративу

Второе возражение связано с контринтуитивностью тезиса о первичности нарратива. Это возражение высказывает, в частности, другой российский аналитический философ, И. Гаспаров. Он считает, что выбор «в пользу нарративного подхода не выглядит очень привлекательным, поскольку, на первый взгляд, автобиографический нарратив кажется вторичным по отношению к метафизическим, моральным и иным убеждениям, которые имплицитно или эксплицитно имеются у его создателя» [Гаспаров 2018, 182]. Гаспаров указывает, что единство конструкции нарратива, его перспектива, траектория и телеологическая направленность определяются лежащими в его основе убеждениями и желаниями личности. Личность первична по отношению к нарративу, и, следовательно, эта же личность могла создать другой нарратив. Августин, к примеру, остался бы Августином, даже если бы не изменил манихейству и создал другую автобиографию. Такую ситуацию Гаспаров считает реалистичной. Как вообще текст может предшествовать личности?

Должен согласиться, что в нашей обыденной практике тексты действительно кажутся вторичными по отношению к личностям. Личности обычно являются авторами и предше-

ствуют написанным ими сочинениям. Но эта ситуация не является единственно возможной. Могут рождаться и тексты, не имеющие автора, и эти тексты могут породить личность в качестве своих героев. Подробно механику этого процесса я попытаюсь раскрыть в конце этой главы. Но здесь в качестве ответа на это возражение приведу один лишь подобный пример сторонника нарративной теории Д. Деннета. Это игра «Психоанализ», описанная в «Объясненном сознании» [Dennett 1991b, 10].

По сути, эта игра представляет собой розыгрыш. Разыгрываемого назначают ведущим и говорят ему, что, когда он покинет комнату, один из участников расскажет остальным свой сон. Сюжет этого сна ведущему надо будет затем восстановить, задавая участникам вопросы, предполагающие только ответы либо «да», либо «нет». И далее ему нужно угадать, кому из компании этот сон приснился. Но как только разыгрываемый покидает комнату, участники договариваются о принципе, по которому будут формироваться ответы. Например, этот принцип может быть таким: если последнее слово в вопросе заканчивается на букву из первой половины алфавита – ответ положительный, в противном случае – отрицательный.

Таким образом, сюжет «сна» не придумывается заранее, но формируется из гипотез ведущего по произвольно установленному алгоритму. При этом те, кто играл в подобную игру, знают, какими увлекательными порой получаются сю-

жеты. В этих сюжетах есть события, герои, развязка... при этом у них в строгом смысле нет автора. Нарратив сам производит своего героя. Почему подобным образом не может складываться автобиографический нарратив?

Возражение 3. Проблема кого-то еще

Нарративный подход предполагает, что сущностью личности является последовательная жизненная история. И что для реализации этой сущности необходима способность структурировать свой опыт, а также способность хотя бы отчасти выражать его: описывать поступки от первого лица и связывать их со своими ментальными состояниями. Но такие способности отсутствуют у людей в младенчестве, а иногда и в старости или в состояниях психических заболеваний (при слабоумии, в хроническом вегетативном состоянии и т. п.). Таким образом, кажется, что нарративный подход вынуждает нас вычеркивать эти этапы из жизни личности. А это противоречит интуициям. И порождает новую проблему – *Проблему кого-то еще*. Если человек в хроническом вегетативном состоянии (ХВС) – не та же личность, что был до этого состояния, то кто этот человек в ХВС? Кто этот «кто-то еще»?

Я могу представить два возможных ответа нарративистов на этот вопрос. То, какой из них правдоподобнее, зависит от наблюдателей и рассказчиков истории жизни агента в ХВС.

Согласно одному ответу, личность в вегетативном состоянии перестает существовать. Вместо нее продолжает существовать только биологическое существо. Это биологическое существо жило и раньше, и личность существовала на его основе. Возможно, оно еще имеет потенциал вновь стать личностью, но в ХВС оно таковой не является.

Согласно другому ответу, личность продолжает существовать. И существование этой личности возможно благодаря историям других рассказчиков, а не рассказам самой личности. Этот новый этап жизни личности можно включить в ее историю как главу о ней, написанную не самой личностью, а другими авторами. Когда речь шла о требовании правдоподобия нарратива, я уже говорил о необходимости согласования нарратива от первого лица с нарративами, выстроенными из других перспектив. Тем самым уже предполагалось, что истории других людей релевантны по отношению к биографическому нарративу личности. Биографический нарратив – это не только история, написанная от первого лица. Даже в условиях полного «молчания» личности он может быть дополнен рассказами других участников и свидетелей событий из жизни личности. Это дает возможность включить в историю личности даже такие этапы ее жизни, как раннее детство или ХВС. Такой вариант нарративного подхода обходит проблему «кого-то еще».

Отмечу, что такая трактовка не нарушает условия личности, о которых говорилось выше. Наличие сознания и само-

сознания, способность занимать личностную установку, владение языком и т. д. – все это, как было отмечено, является лишь нормативными критериями, которые задают идеальную модель личности. Реальные личности удовлетворяют этим условиям лишь отчасти или на протяжении некоторого времени своей жизни. Даже здоровый человек во время сна, очевидно, не отвечает многим из этих критериев. Однако мы не склонны считать, что в это время жизнь его личности прекращается, каким-то образом возрождаясь затем снова.

Аналогичным образом дети, люди с психологическими дефектами или находящиеся в ХВС являются личностями в силу того, что они соответствовали или могут соответствовать этим условиям в будущем, хотя и не отвечают им в данный момент. И именно в свете такого признания их личностями участие других в создании их биографического нарратива оправданно. Напротив, люди с некоторыми тяжелыми врожденными когнитивными дефектами, как и животные или неодушевленные объекты, в принципе не в состоянии удовлетворять этим условиям, и потому они не являются личностями, не имеют и не могут иметь автобиографического нарратива, и другие личности не могут им помочь в его создании.

Возражение 4. Эпизодические личности

Нарративный подход можно критиковать также с помо-

щью контрпримеров. Такую тактику избирает американский философ Г. Стросон. По его мнению, существует тип людей, которые, кажется, не подойдут под критерии личностей, представленные сторонниками нарративного подхода. Таких людей Г. Стросон называет *эпизодическими личностями*. Эпизодические личности – это те, кто не считает себя кем-то, кто существовал в прошлом и будет существовать в будущем. При этом эпизодические личности, естественно, понимают, что биологическое существо, в котором реализуется их Я, обладает временной длительностью, но само Я они не считают длящимся [Strawson 2005, 65]. Эпизодические личности не видят свою жизнь как разворачивающийся нарратив. В этом смысле они живут только в настоящем, полностью в настоящем. Их Стросон противопоставляет *диахроническим личностям*. Диахронические личности осознают себя как того, кто существовал в прошлом и будет существовать в будущем. Возможно, диахронических личностей большинство. Но, по мнению Стросона, эпизодические тоже существуют. К эпизодическим личностям он относит и себя. Как бы то ни было, помыслить существование эпизодических личностей можно. Значит, в наших интуициях обнаруживается противоречие. Из этого, по мнению Стросона, следует, что, согласно интуициям, нарратив не является неотъемлемым и сущностным элементом личности. А значит, нарративный подход ложен.

Стросон предвидит соображения, которые мог бы выдви-

нуть сторонник нарративного подхода в свою защиту, – аргумент, доказывающий нарративный характер структуры даже эпизодических личностей. Аргумент состоит в следующем. Жизнь эпизодических личностей – это *развитие*, как и жизнь диахронических личностей. В этом смысле жизнь эпизодических личностей тоже представляет последовательность. Эту последовательность могут описывать как сами эти личности от первого лица, так и их окружение. Следовательно, эпизодические личности также реализуются в нарративах. Но Стросон считает, что этот аргумент неудачен, так как он сводит нарративный подход к тривиальности. Если жизнь эпизодической личности можно тем не менее считать реализующейся в нарративе, то и жизнь любого человека будет в этом смысле нарративной. Более того, даже жизнь лошади или собаки можно таким образом считать реализующейся в нарративе. В конце концов, и в жизни лошади или собаки будут развитие и последовательность. Таким образом, Стросон предлагает своеобразную стратегию: либо нарративный подход тривиален, либо он опровергается контрпримером. Так ли это?

Мне кажется, что критика Стросона может быть эффективна для некоторых нарративных концепций личности. Но она бьет мимо тезиса, который я пытаюсь защищать в этой работе. Для удобства свой тезис я назову *Прагматическим тезисом* о нарративной природе личности. Суть его в том, что именно моральная ответственность за действия лично-

сти в прошлом (а также другие прагматические вопросы, такие как забота о выживании или справедливой компенсации) базируется на нарративной структуре личности. Прагматический тезис следует отличать от *Психологического тезиса* о нарративной природе личности⁴⁸. Согласно Психологическому тезису, все люди переживают свою жизнь как историю, нарратив. Именно Психологический тезис уязвим для критики с помощью контрпримеров с эпизодическими личностями, а не Прагматический. Прагматический тезис предполагает лишь то, что основанием применения моральной ответственности будет нарративная структура и единство личности. Но человек может существовать и без такой структуры. В таком случае, однако, он просто не будет личностью в том понимании этого слова, как оно используется в данной работе, и применение моральной ответственности к нему будет неуместным, неоправданным, безосновательным. Моральная ответственность – как плод на ветке нарратива: падает ветка, падает и плод. Значит ли это, что эпизодические личности не являются собственно личностями и не несут ответственности? В частности, означает ли это, что к Стросону нельзя применять эту категорию?

В строгом смысле – да. Действительно, эпизодические личности не должны нести ответственность за действия. Но

⁴⁸ Г. Стросон также выделяет *Этический тезис* о нарративной природе личности. Согласно этому тезису, переживание и осознание жизни как нарратива является необходимым для благополучной жизни [Ibid., 63].

является ли личность эпизодической или диахронической – вопрос не такой однозначный. В решении этого вопроса не только сама личность, но и представления окружающих об этой личности имеют значение. Человек может декларировать отсутствие особой привязанности к своему прошлому и особой заботы о своем будущем, как это делает Г. Стросон. «Я имею прошлое, как любой другой человек, и я прекрасно понимаю, что имею прошлое, – пишет он. – У меня достаточно фактических знаний о нем, и я также помню некоторые из моих прошлых переживаний “изнутри”, как говорят философы. И при этом я ни в коей мере не переживаю свою жизнь как оформленный нарратив и тем более как бесформенный нарратив. Совсем не переживаю. Точно так же у меня нет какого-то великого или особого интереса к своему прошлому. Как нет у меня и какой-то серьезной озабоченности своим будущим» [Ibid., 67]. И при этом из его поведения может следовать обратное, по крайней мере иногда. В таком случае его нельзя считать эпизодической личностью. К примеру, если человек все-таки иногда сожалеет или гордится событиями своего прошлого, если он иногда с большей тревогой относится к переживаниям в будущем собственной боли, чем к будущим переживаниям боли других, значит, он – не эпизодическая, а диахроническая личность. И он несет моральную ответственность за действия, которые следует относить к его биографической истории, даже если сам он их явным образом к себе не относит. Каким человеком является Гален

Стросон? Мне сложно сказать, но мне трудно представить, чтобы человек, живущий вполне обыкновенной жизнью, был эпизодической личностью.

С другой стороны, я, конечно же, могу представить себе человека, проживающего экстраординарную жизнь, который был бы эпизодической личностью. Таким человеком мог бы быть, например, отшельник, проводящий время в медитации и реализовавший буддийские идеалы: человек, лишенный страдания и освободившийся от желаний. Нет ничего противоречивого в том, чтобы помыслить человека с такими психологическими характеристиками. Но его поведение, по моему мнению, радикально отличалось бы от поведения обычных людей, и это было бы очевидно со стороны. Такого человека вряд ли можно было бы считать обычной личностью или вообще личностью. Более того, думаю, ему вообще трудно было бы приписывать наличие Я.

Это отчасти подтверждают описания переживаний существования эпизодических личностей самого Г. Стросона: «Я не имею осмысленного ощущения, что я – я, сейчас размышляющий над этим вопросом, – существовал когда-то в прошлом. И для меня очевидно, что это не ошибка восприятия. А скорее регистрация факта обо мне – о том, кто сейчас размышляет над этой проблемой» [Ibid., 68]. Стросон ссылается на то, что у него нет ощущения, что Я присутствовало в прошлом. Но он также не разъясняет, в чем заключается осмысленное ощущение существования Я в настоящем. Если «пе-

реживания перспективы изнутри» не являются достаточными для переживания Я, то ни в прошлом, ни в настоящем с помощью интроспекции нельзя обнаружить Я. Достаточно вспомнить приведенные в начале главы рассуждения Юма на этот счет. Таким образом, в ответ на стратегему Стросона я предлагаю другую: либо большинство людей, считающих себя эпизодическими личностями, фактически являются диахроническими личностями, то есть просто личностями – в том смысле, как этот термин понимается в данной работе, либо такие люди действительно эпизодические личности, но тогда они не только не личности в собственном смысле слова, но и не обладают Я. Эти специальные случаи эпизодических личностей возможны, но они не являются угрозой для Прагматического тезиса о нарративной структуре личности. Таким образом, я считаю критику Стросона неопасной.

Не будучи опасной, она тем не менее небесполезна. Замечания Стросона подтверждают, что свойство «быть личностью» не является фундаментальным и простым свойством объектов. Оно основано на паттерне событий. Как и ментальная каузальность, это – сложное высокоуровневое свойство, различимое только из перспективы моральной ответственности и рождающее иллюзию метафизической простоты, иллюзию точки сборки и источника действий.

Анализ преимуществ нарративного подхода и замечаний в его адрес позволяет сделать вывод о том, что это – наиболее удачное решение проблемы связи событий и атрибуции

их личности. Нарративный подход позволяет объединить атрибуты личности как единомоментно, так и в темпоральной перспективе. Атрибуция событий, в свою очередь, обеспечивает основания приписывания моральной ответственности за действия, произведенные личностью в прошлом. Однако остается не до конца проясненным соотношение нарратива и личности. Ведь именно личность является субъектом ответственности. Нельзя же приписывать ответственность тексту!

Итак, нарратив позволяет охарактеризовать личность. Но кажется, что личность – не то же самое, что биографический нарратив. Представления о нарративе и интуиции о личности, об эго, о Я значительно расходятся. Выше, в начале главы, я уже перечислял традиционные характеристики личности. Среди них – представление, что личность – это простая, элементарная, неделимая сущность. Что она является единым командным пунктом, где принимаются волевые решения. Что эго – «главное место», и что личность сохраняет нумерическое тождество во времени и является абсолютно уникальной, что ее невозможно копировать или размножить. Если нарратив связан с личностью, то, согласно этим первичным интуициям, он каузально зависим от эго, Я. Я является автором нарратива, разве не так?

VI. Личность как центр нарративной гравитации

1. Полезные аналогии, проясняющие, что такое личность

С тем, что субъект, или эго, является подходящим объектом для атрибуции ответственности, сложно не согласиться. Именно эго мыслится рациональным источником воли и «точкой сборки» всех ментальных содержаний. Это – элементарная, неделимая сущность, сохраняющая нумерическую тождественность во времени, уникальная и защищенная от копирования. Но это эго, судя по всему, не является реальным объектом. Это не вещь, не что-то конкретное. Скорее это теоретическая конструкция, абстрактная сущность, в некотором смысле – вымысел. Такую позицию занимают Д. Деннет, Д. Вегнер, М. Газзанига.

Как всякая фикция, эго не может быть каузально эффективным. И потому, согласно нарративному подходу, не абстрактное эго является причиной появления автобиографического текста, а автобиографический текст является источником убеждения в существовании эго с описанными выше характеристиками. В данном случае Я – это не автор, а пер-

сонаж текста. Но не текст создается персонажем, а персонаж – текстом. Представление об этом персонаже и его характеристиках рождается в результате интерпретации поведения организма и произведенного мозгом текста. Интерпретацию осуществляют как другие, так и сам агент. При этом представление об эго – не случайный вымысел. Это очень полезная конструкция, так как она позволяет делать успешные предсказания, влиять на поведение организма и взаимодействовать с ним. Ее онтологический статус не многим хуже, чем статус научных понятий, таких как крутящий момент, электрическая мощность или, например, центр тяжести в физике. Для разъяснения такого представления Деннет предлагает несколько аналогий [Dennett 1992].

Аналогия 1. Центр тяжести. Центр тяжести играет важную роль в определении поведения объекта. Возьмем, к примеру, неваляшку, некогда одну из самых распространенных детских игрушек. Как ни наклонять эту фигурку в разные стороны, она в итоге примет исходное вертикальное положение. Это происходит потому, что ее центр тяжести находится в нижней части игрушки и стремится занять самое нижнее положение. Наклоняя фигурку, мы смещаем эту точку вверх. Но стоит только ее отпустить, как центр тяжести вновь займет самое близкое к опоре положение.

Зная законы физики, можно найти точку, в которой находится центр тяжести любого объекта. Теоретически в некоторых случаях можно указать даже конкретный атом, совпа-

дающий с центром тяжести. Но пространственное совпадение не будет означать тождественности. Недопустимо идентифицировать центр тяжести объекта с конкретным атомом. Это было бы категориальной ошибкой. Атом обладает физическими свойствами: массой, размером, электрическим зарядом, а центр тяжести – нет. Более того, центр тяжести некоторого тела может находиться и вне границ этого тела – как это обычно бывает, например, у кольца. Это просто геометрическая точка в пространстве. И вести себя она может иначе, чем частицы материи. Так, если прикрепить к одной доске другую другой формы, центр тяжести объединенного предмета окажется в другом месте. Но перемещение центра тяжести будет не последовательное, как у всех материальных объектов, а дискретное, скачкообразное. Некоторые физические ограничения не действуют на абстрактные теоретические сущности.

Физик, изучая движение вещей, допускает существование центра тяжести. А антрополог-любитель, каковыми мы все являемся, постулирует существование эго как в себе, так и в других людях. Оба случая принципиально схожи, так как пополняют онтологию теоретическими сущностями, полезными в реализации практических целей. И эти теоретические сущности имеют любопытные особенности. Эти особенности можно исследовать на примере другой фикции – вымышленных персонажей.

Аналогия 2. Персонаж литературного произведения. От-

кроем роман «Моби Дик». На первой странице читаем: «Зовите меня Измаил». Кого называть Измаилом? Автора этих строк, Мелвилла? Нет, мы зовем Измаилом Измаила. Но кто он такой? Это вымышленный персонаж Мелвилла. Довольно много о нем мы узнаем из романа. Что он китобой и что он ищет приключения на судне «Пекод». Кое-что оказывается понятным между строк. Например, что у него две руки и две ноги. Ведь кроме особенностей, описанных автором, мы полагаем, что его герой в остальном похож на прочих людей (в отличие от капитана Ахава, например). Но что-то об Измаиле останется принципиально неопределенным. К примеру, было ли у него родимое пятно на плече? Этот вопрос не может быть выяснен, поскольку Мелвилл об этом ничего не говорит. Существование такого рода неопределенностей и есть основное отличие вымышленных персонажей от исторических. В отношении первых не всегда действует принцип бивалентности. Утверждения о них могут носить не только истинный или ложный характер, но и *неопределенный*. Ведь их существование задается только границами текста и культуры, в которых они создаются. То же самое применимо и к личностям. Как, впрочем, и к центрам тяжести. Они имеют только те свойства, какими наделила их теория. Что, однако, не делает эти понятия бессмысленными или неприемлемыми для серьезной дискуссии и даже науки. В конце концов, наука не смогла бы развиваться без подобных понятий.

История с Измаилом – хорошая иллюстрация появления

вымышленных личностей. Но она, кроме пользы, может отбросить нас назад, вернув к исходным интуициям. Она может напомнить о связи вымышленных героев с реальными авторами. Для нейтрализации этой интуиции можно обратиться к еще одной ситуации, придуманной уже упомянутым Д. Деннетом специально для этой цели, – к ситуации с машиной, которая пишет романы. Это будет третья иллюстрация, которую мы используем для прояснения сущности личности.

Аналогия 3. Созданный алгоритмом образ. Представьте машину-писателя. В ней есть набор различных текстовых фрагментов, правил и каких-то алгоритмов перемешивания слов. Эта машина создана не для написания какого-то конкретного романа, а просто чтобы генерировать осмысленные тексты. Создана так, что ее создатели сами не знают, какой сюжет и каких героев она в этот раз выпишет. Вот мы включаем машину, и из ее принтера появляется заглавие, а за ним – первые строки романа: «Описание жизни Гилберта». «Зовите меня Гилберт», – начинается текст. И далее появляются страницы биографического описания жизни и приключений Гилберта. Машина, не имеющая представлений ни о мире, ни о романе, вполне способна создать и вымышленного героя, и его биографию. Теперь усложним ситуацию: поместим машину в небольшого робота на колесах. Представьте, как он вместе с машиной-писателем попадает в чулан и там случайно закрывается. И тут же из его принтера появляется продолжение истории: «Меня заперли в чулане. Помогите!»

Помочь кому? Гилберту, наверное. Но кто этот Гилберт? Робот? Или вымышленный персонаж, созданный роботом? Мы спасаем машину из чулана, и она тут же выдает новый текст: «Благодарю. Ваш Гилберт». Теперь уже нельзя игнорировать любопытное соответствие между историями вымышленного персонажа, Гилберта, и робота. Но сам робот – не персонаж, он может даже не знать, что создает персонажа. Он фактически не обладает качествами, приписываемыми персонажу, в частности той целостностью, рациональностью и сентиментальностью Гилберта. Робот – только машина с различными механическими частями. И тем не менее эта машина способна породить образ рационального, чувствительного, целостного эго. Скорее всего, машина человеческого мозга делает что-то подобное.

Доводы в пользу такого суждения и релевантности приведенных выше трех аналогий можно почерпнуть, в частности, из области психиатрии и нейронауки. В качестве основы для одного из доводов может служить *феномен множественной личности* (иначе это психическое нарушение называется диссоциативным расстройством личности). Другой довод можно построить на наблюдениях за *пациентами с расщепленными полушариями мозга*.

2. Эмпирические доводы в пользу концепции Центра нарративной гравитации

А. Диссоциативное расстройство личности

Синдром множественной личности (СМЛ) – это психическое расстройство, при котором один человек, один организм обладает двумя или более идентичностями или личностными состояниями. Каждая «личность» (или альтер-эго) имеет собственные характер, мировоззрение и воспоминания (а иногда и собственные возраст, расу, пол), отличающиеся от таковых характеристик других «личностей». Эти «личности» попеременно контролируют поведение человека. Поступки, совершенные под контролем одной «личности», не оставляют в памяти другой «личности» следов. Поэтому как минимум одна «личность» не знает о существовании другой. Свидетельства о существовании подобного синдрома существуют с XVII века, но именно на конец XX в. приходится основная масса исследований этого заболевания. К 1995 г. было зарегистрировано почти 40 000 подобных пациентов. Большая часть из них диагностирована в Северной Америке.

Обычно у расщепленной личности есть «хозяин», домини-

нирующая по времени личность. Эта личность имеет довольно слабую эмоциональность. В отличие от нее альтер-эго демонстрируют доминирование какой-то одной, но сильной эмоции или качества: раздражительности, заботливости, сексуальности, инфантильности и т. п. Жизнь альтер-эго состоит из тех эпизодов, когда оно выходит на сцену. Эти эпизоды оно запоминает и из них составляет историю о том, что оно такое. Чаще всего альтер-эго имеют собственные имена. Количество таких личностей может быть от двух до нескольких десятков. Замечено, что, когда «семья» становится большой, начинают появляться личности противоположного пола. Своеобразие феномена можно почувствовать, только если рассмотреть какой-нибудь конкретный пример. Возьмем историю болезни, воспроизведенную Деннетом в статье «Говоря за нас». Это история о Мэри. История эта вымышленная, но она не так далека от реально зафиксированных историй пациентов с подобными симптомами.

Тридцатилетняя Мэри страдает депрессией, резкой переменой настроения и потерей памяти. Последние несколько лет она часто попадает в клиники. Ей ставят разные диагнозы: шизофрения, маниакально-депрессивный психоз, пограничное расстройство личности. Ее случай наконец привлекает внимание доктора Р., занимающегося диссоциативными расстройствами. Мэри рассказывает врачу свою историю.

Ее отец умер, когда ей было два года. Поэтому Мэри воспитывал отчим. В целом он был добр к Мэри, но иногда «это

заходило слишком далеко». В подростковом возрасте для нее характерны были резкие изменения настроений, плохая дисциплина в школе. Ее воспоминания о детстве, о школе и о занятиях очень отрывочны. К примеру, она хорошо играет на гитаре, но не помнит, как этому научилась. Рассказывая о себе, она иногда начинает говорить в третьем лице. Сейчас она также обнаруживает провалы в памяти: находит в гардеробе одежду, которую не покупала, отправляет родственникам несколько повторных поздравительных открыток. Она утверждает, что у нее высокие моральные ценности, но признаёт, что у других о ней может складываться противоположное мнение.

На одном из сеансов врач решает идти ва-банк. С согласия Мэри он вводит ее в гипнотическое состояние и просит появиться ту часть Мэри, что не выходила раньше. И вдруг Мэри, или та, что недавно была Мэри, неожиданно бросает доктору кокетливый взгляд: «Привет, док! Я Салли... Эта Мэри – просто тряпка. Она думает, будто все знает, но я могу рассказать тебе гораздо больше». Внезапно появившись, Салли так же внезапно исчезает. На следующих сеансах на сцене появляются новые лица – «Хейти» и «Пегги». Мэри об их существовании, конечно, ничего не знает. Ее истории диссоциированы. Чтобы вылечить пациентку, врачу нужно познакомить различные «личности» Мэри друг с другом и интегрировать их в одну, а также выяснить причину возникшего расстройства. Только в таком случае Мэри сможет вер-

нуться к нормальной жизни. По крупице врачу удается восстановить историю заболевания.

С четырех лет отчим стал брать Мэри с собой в кровать. Он ласково называл ее Сандрой, гладил и ласкал ее. Он говорил, что это будет их секретом. Иногда Мэри старалась доставить радость отчиму, иногда притворялась бесчувственной. Когда же, наконец, Мэри сказала ему, что больше не может терпеть и что расскажет обо всем другим – он ударил ее и предупредил, что тогда они оба отправятся в тюрьму. Переживания оказались невыносимыми. Поэтому Мэри просто перестала ассоциировать себя с Сандрой. Через какое-то время и Сандра «расщепилась», уступив место Салли – той части, что старалась доставить отчиму удовольствие, Хейти – той, что испытывала злобу и гнев, и Пегги – той, что притворялась бесчувственной. Во взрослой жизни Мэри эти личности нашли себе подходящие роли. Когда Мэри испытывала злобу, ее место занимала Хейти, когда ее целовал мужчина, на сцене возникала Салли. В юности все эти перемены объяснялись окружением Мэри как особенности «трудного возраста». К тридцати они уже доставляли ей серьезные проблемы.

История Мэри вполне типична для пациентов с СМЛ: детская травма и внутренний конфликт, затем – неосознанная выработка внутреннего решения и изоляция травмирующего опыта, структурные изменения личности и терапия, вследствие которой травмирующий опыт восстанавливается

и нейтрализуется. Эта история демонстрирует работу одного из защитных механизмов психики – диссоциации. Суть этого механизма заключается в том, что человек перестает ассоциировать действующего персонажа собственных воспоминаний с самим собой и отдает эту роль «кому-то другому», перенося на него и связанные с этими воспоминаниями переживания. Такая способность может реализовываться и у здоровых людей в измененных состояниях сознания или в критические, особенно стрессовые моменты. Она представляет собой любопытный феномен.

Но история Мэри в частности и феномен СМЛ в целом для нашего исследования полезны не столько как иллюстрация диссоциации, сколько как модель обратного, более базового, стандартного процесса – *процесса ассоциации или индивидуации личности*, выработки собственного Я. Люди, страдающие СМЛ, ассоциируют ментальные содержания во множество кластеров, а здоровые – преимущественно в один. Если нарратив последователен, то в процессе индивидуации появляется один агент. А если нарратив не может быть выстроен последовательно ввиду его травматичности – то много. Но суть процесса от этого не меняется. Как пишет гарвардский психолог Д. Вегнер, субъективное эго, которое есть у нас в таком контексте, – это только один виртуальный агент из множества возможных [Wegner 2002, 255]. И, судя по описанным случаям, процесс индивидуации личности, появления этого виртуального агента, больше похож на выработ-

ку писателем образа персонажа или постулирование теоретической сущности, чем на обнаружение нового природного объекта.

Для взаимодействия с внешним миром и предсказания даже собственных действий удобно постулирование единой, рациональной, уникальной, непрерывно существующей личности. Но если сделать это не получается, один организм, один мозг может создать два образа или даже целую когорту. Личность способна интегрироваться или дезинтегрироваться, как центр тяжести: стоит разделить физический объект, и у его частей тут же возникнут два собственных центра. Но если мы объединим части вновь, центр тяжести снова будет один. Какой центр тяжести был новым, какой – старым? Существовали ли эти центры одновременно в одном объекте? Будет ли вновь объединенный центр тяжести тем же самым, что был раньше? Такие вопросы можно задавать. Но если ответы на них и будут различаться, это будут, скорее всего, лишь терминологические различия. То же самое верно и в отношении вопроса о количестве и тождестве личностей Мэри. Можно сказать, что в действительности у нее одна личность, но она многогранна, а можно сказать, что у нее 4 личности. Оба ответа будут в какой-то степени верны и будут различаться лишь терминологически. Если личность, как следует из наших размышлений, это абстрактная сущность, данные вопросы не имеют принципиального значения.

В. Пациенты с разделенными полушариями

Представление о том, что личность как единый рациональный субъект представляет собой удобную фикцию, поддерживается также наблюдениями за людьми с расщепленными полушариями мозга (*split-brain patients*). Эти люди в лабораторных условиях демонстрируют удивительный феномен – раздельную, автономную интеллектуальную работу полушарий мозга. Исходя из этих наблюдений, можно допустить у них два раздельных потока сознания с различными воспоминаниями, желаниями, убеждениями. При этом в обычных условиях такие пациенты практически неотличимы от здоровых людей. Ни люди со стороны, ни даже сам субъект не обнаруживает аномалию. Кажется, что даже в патологических случаях психика находит способы ассоциации переживаний и ретроспективного объяснения поведения, как будто оно происходит из единого волевого центра. Организм домысливает существование интегрированного агента даже тогда, когда он очевидно отсутствует.

Первооткрывателями феномена автономной работы полушарий были нейропсихологи М. Газзанига, Р. Сперри и Дж. Боген. С начала 60-х они наблюдали за поведением пациентов, которые прошли операцию по рассечению мозолистого тела, сплетения нервных волокон, образующих прямую связь между правым и левым полушариями. Эта опера-

ция называется каллозотомией. Каллозотомия использовалась в качестве средства борьбы с тяжелыми формами эпилепсии. Она позволяла локализовать эпилептические симптомы в одной части мозга. Таким образом, в результате каллозотомии припадки эпилепсии обычно пропадали, а пациент после нескольких недель восстанавливался. Но восстановление, как выяснили Газзанига, Сперри и Боген, происходило только «с виду». Пациенты никогда не возвращались к исходному психологическому состоянию. Эксперименты Газзаниги и его коллег показали, что после операции полушария мозга начинали работать как два изолированных, автономных рациональных модуля. Оба полушария оказываются способны раздельно воспринимать и анализировать информацию, самостоятельно инициировать поведение. Между ними возможны разногласия и конфликты. Как это ни удивительно, сам пациент при этом как будто ничего не замечает. Не замечали раньше эти внутренние конфликты и исследователи⁴⁹. Как такая серьезная дезинтеграция внутренних процессов могла оставаться незамеченной?

Избежать очевидных нестыковок и конфликтов помогала ориентация на косвенные внешние сигналы. Разделенные полушария мозга, не имея прямого канала сообщения, использовали косвенные сигналы, чтобы «общаться» друг с

⁴⁹ Верификацию полного восстановления функциональности мозга провел на 26 пациентах с рассеченным мозолистым телом Э. Акелайтис. Ему не удалось найти никаких особенностей в постоперационном поведении этих пациентов [Akelaitis 1945].

другом. Они ориентировались на внешние проявления работы соседнего полушария, такие как движения противоположной части тела или речь. Полушария вели себя как два различных, но близких человека, договорившиеся о полном согласии во всем: когда один что-то говорил (а поскольку речевые центры располагаются, как правило, только в левом полушарии, правое не может контролировать речь и участвовать в вербальных отчетах), другой внимательно слушал, когда один смеялся, другой делал вид, что тоже обнаруживал что-то смешное. При этом субъект вовсе не обращал на имевшуюся дезинтеграцию внимания и вел себя так, как будто после операции в его внутренней жизни ничего не изменилось. Обнаружить патологию удалось только с помощью хитроумных экспериментов, которые и провели нейропсихологи. Один из экспериментов был построен следующим образом.

Известно, что на левых половинах сетчаток каждого глаза (воспринимающих, соответственно, объекты из правой половины зрительного поля) есть области, передающие информацию только левому полушарию головного мозга. На правых же половинах сетчаток (воспринимающих объекты слева) есть области, связанные только лишь с правым полушарием. В описываемом эксперименте пациенту с разделенными полушариями демонстрируют две картинки, расположенные так, чтобы каждое из полушарий получало данные лишь об одной из них. К примеру, изображение лапы цыпленка

располагают справа под таким углом, чтобы информация о нем обрабатывалась только левым полушарием, а картинку с заснеженным пейзажем размещают слева так, что его «видит» только правое полушарие. После этого пациенту предлагают выбрать из набора других картинок те, которые ассоциативно связаны с начальным образом.

Пациент правой рукой выбирает цыпленка, а левой – лопату. (Лопата в данном случае – правильная ассоциация со снежным пейзажем.) Руководитель эксперимента просит пациента обосновать свой выбор. Пациент отвечает: «Все очень просто. Цыпленок ассоциируется с лапой цыпленка. А лопата нужна, чтобы чистить курятник». Интересно, что левое полушарие, обнаруживая поведение левой руки (управляемой правым полушарием), пытается интерпретировать его, используя ту информацию, которая в ходе эксперимента была доступна именно ему, левому полушарию, не используя информацию, которая была доступна правому, и даже не предполагая, что реакция левой руки могла быть обусловлена чем-то другим. Учитывая, что пациент осведомлен о прошедшей операции и имеет достаточное представление о ее последствиях, гораздо логичней был бы другой ответ: «Я не знаю причины выбора лопаты. Возможно, какая-то другая информация была передана моему второму, изолированному полушарию». Удивительно, но никто из пациентов не приходит к такому выводу. Они даже не обнаруживают аномалию. Этот феномен стабильно проявляется и в других экс-

периментах с пациентами с разделенными полушариями.

В одном из них какую-нибудь команду, например команду «улыбайся», предъявляют из левой области визуального поля правому полушарию пациента. Пациент улыбается. Руководитель эксперимента просит объяснить причину улыбки. Пациент объясняет: «Вы, ребята, приходите и тестируете нас каждый месяц. Вот классная работка!» Если показать правому полушарию пациента команду «встань», пациент встанет и пойдет. Когда его спросят почему, он объяснит: «Хочу пойти в соседнюю комнату, взять колы». Подобные эксперименты всегда приводят к аналогичным результатам. Что это означает? Как это соотносится с приведенными выше рассуждениями о личности?

Пациенты с разделенными полушариями демонстрируют удивительную черту – способность присваивать и домысливать ментальные содержания. Они делают это абсолютно искренне, незаметно для себя. Они интерпретируют свое поведение как поведение единого рационального агента, при этом таковым, очевидно, не являясь. Они присваивают источник и причину поведения, которая абсолютно не прозрачна для них. Значит, у пациентов встроены соответствующие психические механизмы. Но если такие механизмы есть у пациентов с разделенными полушариями, эти механизмы должны присутствовать и у нормальных, здоровых людей. Тогда, вполне возможно, здоровые люди точно так же ассоциируют ментальные содержания и домысливают суще-

ствование единого длящегося рационального агента. Эти механизмы интерпретируют работу разделенных, автономных когнитивных модулей здорового человека как работу одного целостного рационального субъекта и создают удобную, но фиктивную сущность – личность, главного героя нарратива. Таким образом люди решают практическую задачу идентификации для прогнозирования собственного и чужого поведения и кооперации с другими людьми.

Исследования пациентов с разделенными полушариями мозга и пациентов с диссоциативным расстройством личности позволяют прояснить механизмы интеграции личности, которые имеются у здоровых людей. Эти механизмы предполагают интерпретацию поведения, домысливание ментальных содержаний и ассоциацию их в кластеры, а также формирование из этих кластеров представления об идентичности личности. Ассоциация ментальных содержаний и формирование личности имеет сходство с постулированием теоретических ненаблюдаемых сущностей в естественных науках и созданием персонажей художественных произведений. Но отличается от них в том, что изобретателем личности является не рациональный агент, а система из связанных между собой когнитивных модулей мозга. Представить такое полифоническое решение поможет четвертая аналогия.

Аналогия 4. Голос Америки. Представьте радиостанцию, которая вещает новости на разных языках. Множество журналистов собирают информацию по всему миру. Сотни вне-

штатных и штатных корреспондентов на местах интервьюируют экспертов, аналитиков, очевидцев событий. Собранную информацию они отправляют в офисы компании. Там десятки редакторов обрабатывают полученные сведения. Они сверяют полученные данные так, чтобы информация из одних источников подтверждалась информацией из других, уточняют детали, компонуют информацию по категориям, готовят тексты. Десятки радиоведущих озвучивают новостные сообщения, подавая их с соответствующим тоном и эмоциональной окраской. Звукорежиссеры и другие технические специалисты записывают аудиосообщения и транслируют их слушателям. Менеджеры обеспечивают работу всей инфраструктуры и задают основные направления политики радиостанции. Пусть весь этот комплекс называется «Голосом Америки».

У этого «Голоса» есть многие характеристики личности. Можно сказать, что у радиостанции есть память, чувства, предпочтения, надежды, таланты. «Голос Америки» тяготеет воспоминаниями о вьетнамской войне, не любит коммунизм, но любит яблочный пирог, немного сентиментален. Это правда. Но справедливость всех этих характеристик не означает, что они персонифицированы в ком-то одном из членов команды «Голоса». Эти характеристики сближают всех, кто создает радиопрограмму. При этом ни один человек не знает всего того, что «знает» «Голос Америки», и ни один человек, наверное, не согласен со всем тем, что говорит

«Голос Америки». Это просто условная линия консенсуса как для внешнего мира, так и для составных частей системы. Нечто подобное верно и в отношении личности. Личность представляет собой условную линию консенсуса для связанных, но автономных когнитивных процессов и модулей мозга. Ей приписываются собирательные общие характеристики, которыми не обладает ни один из процессов и модулей в отдельности⁵⁰.

С. Вымышленные мотивы и вымышленные агенты

Выше я показал, что психические механизмы ассоциации позволяют вполне произвольно объединять ментальные содержания в кластеры. Тем самым я стремился доказать, что определение личности – это не *fait accompli*, а непрекращающийся творческий процесс. Выше я также указал на существование механизмов ретроспективной атрибуции мотивов действий, домысливание ментальных содержаний. Так я хо-

⁵⁰ Эта аналогия в том числе оправдывает вывод С. Левина о том, что, согласно нарративному подходу, не только люди могут быть личностями, но и корпорации. В статье «Моральная ответственность и личность корпораций: нарративный подход» он пишет: «Приведенные рассуждения призваны показать, как нарративный подход к тождеству личности может снять одно из самых существенных возражений против корпоративной моральной ответственности. Возражение строится на том, что только личности могут быть морально ответственны, а у корпораций личности нет, и поэтому они не могут быть моральными субъектами» [Левин 2016b, 74].

тел продемонстрировать, что многие атрибуты личности являются фантастическими. Но все эти выводы были основаны на аномальных случаях, на наблюдениях за пациентами с психическими дефектами. Можно ли делать экстраполяцию этих выводов на здоровых людей? Можно ли делать выводы о работе нормальной психики на основе наблюдений за ее нарушениями?

По моему мнению, нет оснований считать, что обнаруженные психические механизмы являются следствиями нарушений. Просто дефекты позволяют явным образом обнаружить механизмы функционирования нормальной психики. И тем не менее я согласен, что справедливость этих выводов для здоровой психики стоило бы подкрепить отдельно. И такая возможность существует. Это можно сделать с помощью примеров, которые приводит уже упомянутый выше профессор психологии Д. Вегнер в книге «Иллюзия свободной воли». Его примеры позволяют подтвердить, что ретроспективная атрибуция ментальных содержаний и их произвольная кластеризация являются стандартными механизмами здоровой психики.

а) Ретроспективная атрибуция ментальных содержаний

«Голодная Лиса пробралась в сад и на высокой ветке увидела сочную гроздь винограда. “Этого-то мне и надобно!” – воскликнула она, разбежалась и прыгнула один раз, другой, третий... но все бесполезно – до винограда никак не до-

браться. “Ах, так я и знала, зелен он еще!” – фыркнула Лиса себе в оправдание и заспешила прочь». Эта басня Эзопа отражает вполне типичную ситуацию: ретроспективную подмену намерений. Исходным мотивом персонажа басни была добыча лакомства, но, когда виноград оказался недоступным, лиса изменила свое отношение. Мотив достать «сочный виноград» сменило решение игнорировать «незрелый фрукт». Подобную подмену совершают и многие люди в повседневной жизни. Часто действие оказывается противоречащим исходным интенциям, а результат – ожиданиям. Но вместо того, чтобы признать расхождение, обнаружить противоречие и удивиться, человек просто меняет исходную диспозицию: «Что? Да именно это я и собирался сделать!» Известно, что актеры, сыгравшие роль литературного персонажа, склонны в дальнейшем разделять его позицию, даже если она противоречила их собственным изначальным установкам. А авторы эссе на заданную тему имеют склонность защищать основной тезис, даже если этот тезис им вначале казался ложным. Выходит, суждения личности о собственных предпочтениях и мотивах отчасти опираются скорее на объективные наблюдения за собственным поведением, чем на интроспекцию. Самыми яркими примерами таких ретроспективных интерпретаций собственных мотивов являются примеры с *постгипнотическим внушением*.

Люди, которые получили инструкции во время гипноза, после пробуждения часто выполняют эти инструкции. При

этом они не помнят внушения, не знают настоящей причины собственных действий и придумывают самые невообразимые мотивы, чтобы объяснить свое неожиданное поведение. Во время одного из сеансов гипнотизер внушил женщине, что она должна взять книгу со стола и положить ее на полку. Пробудившись, она так и сделала. Когда женщину попросили объяснить ее действия, она сказала: «Не люблю я смотреть на разбросанные вещи; место для книг – на полке. Поэтому я и положила ее туда» [Wegner 2002, 149]. Можно предположить, что это стечение обстоятельств и что у женщины действительно были подобные желания еще до гипноза. Что ж, другой случай не оставляет сомнений в творческой природе ретроспективных объяснений. Гипнотизер внушает человеку абсурдное действие: он должен взять с окна цветочный горшок, обернуть его в ткань, положить на диван и поклониться горшку три раза. После пробуждения человек выполняет данную инструкцию. Какое же объяснение он может дать столь странному поведению? «Вы знаете, когда я очнулся и увидел, что там стоит цветочный горшок, я подумал, что, поскольку здесь довольно холодно, этот горшок лучше слегка согреть, иначе цветок погибнет. Поэтому я обернул его тканью. И потом я подумал, что, поскольку диван стоит как раз возле камина, хорошо бы положить цветок греться туда. И я поклонился потому, что остался доволен тем, насколько хорошая идея пришла мне в голову» [Ibid., 150]. К удивлению окружающих, сам человек не счел свои действия

безумными, а определил их как вполне рациональные, ведь он смог дать им объяснение. Мне кажется, эти случаи дополняют аргумент в пользу творческого ретроспективного редактирования мотивов.

Аналогичным образом, пожалуй, можно дополнить аргумент в пользу произвольной кластеризации ментальных содержаний.

*б) Атрибуция ментальных содержаний
виртуальным агентам*

Американский актер и радиоведущий Эдгар Берген стал популярным в 30–40-х гг. в первую очередь как чревовещатель с куклой Чарли Маккарти. Случай из его жизни позволяет предположить, что он настолько сжился с образом куклы, что искренне приписывал ей часть своих ментальных содержаний. Однажды гость зашел в комнату к Бергену и обнаружил, что тот не репетирует, а самым настоящим образом разговаривает со своей куклой. Берген задавал Чарли философские вопросы, а Чарли отвечал с сократической мудростью. Увидев гостя, Берген смутился, покраснел и сказал, что у него сейчас состоялась удивительно интересная беседа. На что гость заметил, что Берген вел эту беседу сам с собой. «Наверное, в конечном счете так и было, – ответил Берген, – но я задаю Чарли вопросы, а он отвечает. У меня нет ни малейшей догадки, что он скажет в следующий момент, и я по-

ражен его пронизательностью»⁵¹. Возможно, этот случай – голливудский миф. Но мне кажется, что он вполне реалистичен, учитывая творческий талант и развитое воображение актера. Но если читателю трудно поверить в его правдоподобие или если этот пример кажется ему признаком душевной болезни, стоит просто представить на месте Э. Бергена ребенка. Тогда такое поведение должно перестать быть аномалией. Дети часто воображают подобных собеседников. И в современной психологии это считается признаком становления здоровой психики.

Исследования показывают, что примерно 65 % детей имеют воображаемых друзей [Klausen, Passman 2006, 349]. Воображаемый друг – это существо, придуманное ребенком, с которым ребенок регулярно общается. Такими друзьями могут быть невидимые существа, одушевленные в воображении ребенка предметы (в том числе игрушки), а также вымышленные агенты, роль которых разыгрывает сам ребенок. Систематические исследования этого феномена психологами начались в конце XX в. И теперь есть даже основания говорить о появлении нового поля научных исследований [Klausen, Passman 2006].

Изначально наличие вымышленных друзей исследователи интерпретировали как отклонение от нормы. Одни психологи считали, что их присутствие свидетельствует о патологической неспособности ребенка отличать вымысел от ре-

⁵¹ Случай пересказан по [Ibid., 221].

альности, иногда – о психической болезни или неблагоприятных обстоятельствах воспитания. Другие, напротив, полагают, что этот феномен свидетельствует о необычайной одаренности детей, служит показателем особых творческих способностей. Одним из первых, кто считал наличие вымышленных друзей этапом нормального психического развития, был швейцарский психолог Ж. Пиаже. Он сделал вывод, что общение с вымышленными друзьями служит транзитной стадией развития языковых способностей (от внутренней речи – *self-talk* – к социальным коммуникациям). Впоследствии мнение о том, что наличие в раннем возрасте воображаемых друзей является нормой развития личности, закрепилось. В частности, его придерживается М. Тейлор, профессор психологии из Университета Орегона, посвятившая феномену воображаемых друзей монографию [Taylor 1999]. Она приводит следующие случаи в качестве примеров работы здоровой детской психики.

Одна 4-летняя девочка, участвовавшая в исследовании, рассказала о том, что дружит с двумя невидимыми птичками по имени Натси и Натси, которые живут на дереве за окном ее спальни. По словам девочки, у двух Натси цветные перья, они 30 см ростом и постоянно болтают. О двух птичках знали даже родители. Более того, Нат-си были фактически домашними питомцами. Когда семья выезжала куда-то, птички ехали на крыше машины; когда все садились за стол, птичкам оставляли место. И все вместе смеялись их глупым вы-

ходкам. Спустя два года девочка все еще помнила про своих питомцев, и ее мама рассказывала, что они иногда садились с дочкой и вспоминали птичек Натси и Натси [Ibid., 8–9].

Такое описание является характерным для случаев с вымышленными друзьями. Однако существует множество вариаций. Вымышленные персонажи могут принимать самые различные формы и воплощения. У некоторых детей есть вымышленные копии их реальных друзей, у некоторых – это копия персонажа из фильма, другие одушевляют мягкие игрушки и даже обычные предметы. Так, у одного мальчика была длительная и бурная история отношений с комодом. Другие же «разыгрывают» вымышленных персонажей сами. Мама одной девочки сообщила, что ее дочка дружит с вымышленным существом по имени Applejack (Яблочный Пирожок). Когда психологи стали спрашивать девочку, как часто она с ним играет, та стала недовольно возражать: «Это я, я и есть Applejack». Оказывается, мама не разобралась, что девочка не столько дружила с вымышленным героем, сколько иногда им была [Ibid., 15].

Наличие вымышленных друзей в детстве – это распространенное явление и признак нормального развития психики. Ни большое количество таких воображаемых друзей, ни многообразие их форм не являются симптомом отклонений. И существование таких «друзей» во взрослом возрасте также не обязательно является аномалией. Это подтверждает другое исследование М. Тейлор, посвященное отноше-

ниям между авторами художественных произведений и их персонажами [Taylor, Hodges, Kohanyi 2002–2003]. По мнению Тейлор, их отношения во многом схожи с отношениями между детьми и их воображаемыми друзьями.

Ключевое сходство в том, что воображаемые персонажи представляются в обоих случаях как автономные, самостоятельные агенты с собственными мыслями, чувствами и действиями. Эта их автономия выражается в непредсказуемости и неподконтрольности человеку их поведения. Дети часто жалуются, что их воображаемые друзья не ведут себя так, как они хотели бы: не делятся тем, что у них есть, не всегда появляются, когда им бы этого хотелось, или, наоборот, надоедают своим присутствием, разговаривают слишком громко, мешают. На подобную автономию вымышленных персонажей указывают также писатели. В интервью национальному общественному радио Джоан Роулинг, автор мирового бестселлера про Гарри Поттера, на вопрос о том, почему главный герой ее произведения мальчик, а не девочка, отвечала: «Спустя шесть месяцев после начала работы над романом я подумала, что я женщина, а он – мальчик. Но было уже слишком поздно, было слишком поздно делать из Гарри Генриетту. Он был для меня реален как мальчик, и, если бы я надела на него платье, это все равно был бы Гарри, только в женской одежде... Когда я пишу, я никогда не говорю, мол, “так, теперь мне нужен какой-то персонаж”. Как мои герои являются мне – это такой таинственный процесс,

который никто толком не понимает. Они просто появляются» [цит. по: Ibid., 363].

Некоторые авторы «живут» со своими персонажами, беседуют, спорят. Иногда авторы даже вынуждены вести переговоры с героями об их судьбе в романе. Сара Парецки в интервью радио сообщала, что ей пришлось идти на уступки, когда по сюжету ее герой должен был попасть в тюрьму. Герой отказывался идти на это до тех пор, пока она не пообещала ему взамен настоящую любовь [Ibid., 363]. Подобные комментарии, признания наличия независимой и «свободной» воли у персонажей, встречаются у множества писателей и сценаристов. Среди них – Достоевский, Сартр, Пруст, Джеймс, Тарантино. Английский романист Э. М. Форстер так описывает процесс создания романа: «Герои появляются тогда, когда их пробуждаешь, но полные бунтовского духа. Из-за того, что у них так много общего с людьми, они пытаются, как мы, прожить собственную жизнь, и потому часто вступают в заговор против основного сюжета книги. Они “сбегают”, “выходят из-под контроля”: они – творения внутри творения, и часто не помещаются в него. Если им дать полную свободу, они раздерут книгу на куски, но если слишком сильно держать их в узде, они отомстят своей смертью и отравят ее своим гниением» [цит. по: Ibid., 364].

Описание Форстера показывает устойчивость иллюзии независимого существования персонажей. И является хорошим подтверждением нашей гипотезы о том, что здоро-

вая психика может компоновать, кластеризовать ментальные содержания так, чтобы они образовывали нарративы. Если этот нарратив один, в результате его интерпретации появится единственный агент, если множество – агентов тоже будет множество. В любом случае за процессом индивидуации личности стоит творческий процесс. Приведенные примеры должны показать, что аналогии Деннета, представленные выше, релевантны по отношению к личности. Они подкрепляют гипотезу о том, что личность – это абстрактная сущность, некоторая удобная фикция, рождающаяся в ходе интерпретации нарратива. Эта гипотеза, в свою очередь, позволяет нам разрешить проблему, с которой мы начинали эту главу, – Проблему исчезающей личности.

Впрочем, можно было бы попытаться возразить: если личность – это фикция, и при этом, как утверждалось выше, личность необходима для реализации свободы воли и моральной ответственности, значит, свобода воли и моральная ответственность – тоже фикции? Это очень важное замечание, и я должен в заключение представить на него ответ. Да, я действительно считаю свободу воли и моральную ответственность в некотором смысле фикциями, но только в том смысле, в каком их обычно отстаивают сторонники субстанционального дуализма или агент-каузального либертарианства. Свобода воли и моральная ответственность, необходимые для рационализации существующих межличностных отношений, существуют реально. Референтами этих поня-

тий служат объективно существующие абстрактные паттерны, которые, в свою очередь, являются неотъемлемой частью описания реальности. Свобода воли, моральная ответственность и личность – это фикции, но *полезные* фикции, место которым найдется в науках так же, как и понятиям «центр тяжести», «вектор движения» или «ген». Такой онтологии придерживается и Деннет, называя ее «умеренным реализмом» [Dennett 1991a, 30]. Однако наиболее емкое и запоминающееся название этой онтологической позиции дал профессор философии Кейптаунского университета Дон Росс. Он назвал ее «реализмом тропических лесов» [Ross 2000, 147].

VII. Итоги третьей главы

Событийный каузализм создает опасность для тождества личности и самого ее существования. Ее можно назвать Проблемой исчезающей личности. По мнению множества философов, существование и тождество личности являются условиями моральной ответственности. Но если событийный каузализм верен, то эти условия не соблюдаются, личность как бы «исчезает», и некому оказывается приписывать действие и ответственность за него. Препятствием для признания наличия свободы воли и ответственности также является противоречие между обыденными интуициями о свойствах личности и эмпирическими данными о том, как устроен человек и его мозг. Это препятствие можно обозначить, вслед за У. Селларсом, проблемой согласования «данности и научных взглядов» [Sellars 1991] в отношении личности и ее действий. Эта проблема была обнаружена уже философами Нового времени, на нее указывали Г. Лейбниц и Д. Юм. Современная разработка этой проблемы обнаруживается, в частности, у Д. Деннета. Он делает выводы о том, что рациональность, сознательность, самосознание и другие классические условия существования личности не могут быть полностью удовлетворены в принципе, а личности как конкретной, единой сущности не существует.

В поисках решений этих проблем мы обратились к совре-

менным теориям личности и рассмотрели психологический подход Дж. Перри и С. Шумейкера, биологический подход Э. Олсона и субстанциональный подход Р. Суинбёрна. Я попытался показать, что эти подходы оказываются неприменимы к некоторым гипотетическим ситуациям, таким как «Теле-транспортер» Д. Парфита и «Где Я?» Д. Деннета. Эти гипотетические ситуации выявляют проблемы концептуальных построений современных аналитических философов. Возражения к описанным подходам были резюмированы в Аргументе пяти трансплантаций. В контексте этого аргумента и представленных мысленных экспериментов вновь встал вопрос о наличии свободы и основаниях приписывания ответственности.

Нарративный подход в этом отношении представляется наиболее перспективным. Этот подход не является классическим, так как он заменяет вопрос о реидентификации вопросом о характеристике личности и опирается на нереалистическую концепцию личности. Но он позволяет решить проблему рациональных оснований ответственности. Согласно разработанной концепции, каркасом личности является биологический организм. Одна из функций организма – создание расширенного фенотипа. Расширенным фенотипом человека служат в том числе созданные им тексты. Именно эти тексты обеспечивают связь атрибутов человека в прошлом и настоящем, его ментальных содержаний и действий. Таким образом, сама структура текста позволяет из-

бежать вопроса о тождестве личности во времени и заменить его вопросом о принадлежности. Она позволяет снять требование нумерического тождества во времени и заменить его требованием связности, выполнение которого успешно обеспечивается текстом. Структура текста и его смысловая направленность позволяют интерпретировать события, описываемые в нем, как результат деятельности какого-то определенного агента.

Такой агент – это допущение, результат интерпретации, виртуальный, теоретический конструкт. Этот агент является абстрактной, а не конкретной сущностью. Поэтому найти его ни интроспективно, ни объективными методами не удастся. Личность как фиктивная сущность наделяется фантазийными характеристиками, такими как полная рациональность, целостность, неделимость, уникальность. Эти характеристики в абсолютном смысле недостижимы. И личность всегда обладает ими только «отчасти».

В нарративной концепции личность – это фантазия, порожденная структурой текста. Но это – полезная, удобная фантазия. Она позволяет объяснить и предсказать поведение биологического организма и наладить сотрудничество между людьми. Ее удобством обуславливается повсеместное принятие этой теоретической конструкции. А повсеместное принятие является основанием для возможности ответных реакций и моральной ответственности.

Заключение

Моя книга – исследование свободы воли и моральной ответственности. Она отвечает на вопрос, возможны ли ответственность и свобода в мире, где действуют физические законы, где длинная цепочка причин определяет наличное состояние вещей и где причинами могут быть только события, а не субстанции. В решении этого вопроса нельзя было обойтись без современных представлений из области физики, нейрофизиологии, психологии. В этом смысле работа опирается на научные теории. Однако эта книга – *философская*. Сами по себе научные теории и эмпирические факты не могут дать искомого ответа: они не способны склонить ни в ту ни в другую сторону. Только в философском контексте эти научные теории могут свидетельствовать против возможности свободы и ответственности. Поэтому поиск ответа я осуществлял с помощью философских методов: концептуального анализа, мысленных экспериментов и логических аргументов. Так я попытался отвести основные опасности.

Традиционно опасностью для свободы и ответственности считается детерминизм – представление о том, что в каждый момент времени существует только одно реально возможное будущее. Однако в книге я представил и другие, не менее опасные для свободы и ответственности концепции: физикализм и событийный каузализм. Физикализм – это убежде-

ние в том, что у всех событий в мире есть только физические причины, а событийный каузализм – тезис о том, что причинами событий могут быть только другие события. Детерминизм исключает альтернативные возможности, физикализм – ментальную каузальность, а событийный каузализм – связь личности и действия. Три эти философские концепции, как может показаться, подрывают условия для реализации свободы и ответственности. Анализу каждой из этих «опасных» теорий и их соотношения с условиями свободы я посвятил отдельную главу.

Первая глава была посвящена совместимости *ментальной каузальности и физикализма*. Хотя этот вопрос обычно считается частью проблемы сознания, я включил его в дискуссию о свободе воли и ответственности. Это включение оправданно потому, что каузальная эффективность ментальных состояний, желаний и намерений, является определяющим элементом действия, а без возможности действия, очевидно, невозможно и свободное действие. Но ментальная каузальность, возможно, не согласуется с физикализмом. По крайней мере, это следует из Аргумента исключения. В качестве возможных решений я представил несколько философских теорий. Самой удачной из них, как выяснилось, является телеофункционализм. Эта теория наиболее успешно опровергает Аргумент исключения и обосновывает *возможность ментальной каузальности*. В поддержку этой теории я предложил собственный мысленный эксперимент и ком-

ментарии к нему. Новый мысленный эксперимент, как и эксперименты, предложенные теоретиком телеофункционализма Деннетом, являются контрпримерами к Аргументу исключения. С помощью этих контрпримеров первое препятствие на пути свободы было преодолено.

Во второй главе я исследовал отношение между *свободой и детерминизмом*. В начале этого исследования я определил общие *условия свободы воли и моральной ответственности* – *принципы автономии и альтернативных возможностей*. Эти принципы, по мнению некоторых философов, несовместимы с *детерминизмом и даже с индетерминизмом*. Я рассказал о нескольких инкомпатибилистских теориях, которые построены на этой интуиции. Потом я изложил аргументы, которые используются сторонниками этих позиций. Самый известный из них – Аргумент последствий – якобы показывает, что никто в принципе не имеет альтернатив совершаемым действиям, так как будущее целиком является последствием прошлого и законов природы. Но, проанализировав этот аргумент, я показал, что оба принципа могут быть выполнены в условиях детерминизма. Аргумент построен на неоправданно узкой трактовке выражения «иметь возможность». Из моих размышлений следует, что ни детерминизм, ни индетерминизм не являются препятствиями для свободы. А значит, наиболее достоверной позицией является *суперкомпатибилизм*. Согласно этой позиции, какими бы законами наш мир ни определялся, детерминистическими или

индетерминистическими, свобода в принципе реализуема... только если ее не ограничивает событийный каузализм. Ему я посвятил третью главу.

Событийный каузализм – теория, согласно которой причиной действий могут быть только события – является препятствием для *асинхронного тождества личности* (то есть тождества личности в прошлом и настоящем), но он может оказаться препятствием и для возможности приписывать решения и ответственность агенту. Если причинами событий могут быть только другие события, как можно связать личность и ее поступки в прошлом? Этот вопрос связывает проблему тождества личности с проблемой ответственности. В качестве возможных решений я исследовал наиболее авторитетные философские теории тождества личности. Самой удачной из них, как было показано, является нарративный подход. Это антиреалистическая концепция тождества личности. В ней длительное существование личности признаётся зависимым от убеждений других личностей. Эта теория подтверждается экспериментальными данными (в частности, исследованиями Д. Вегнера и М. Газзаниги) и позволяет приписывать решения и ответственность агенту.

Таким образом, в результате исследования я доказал, что ни одна из перечисленных выше философских концепций не исключает свободу и ответственность, то есть условия свободы и ответственности в принципе реализуемы. Конечно, для того, чтобы они фактически реализовались, должны вы-

полняться и другие условия: личность должна быть психологически здоровой и находиться в состоянии бодрствования и здравомыслия, она должна фактически знать о многих обстоятельствах своего поступка, быть в состоянии предвидеть последствия, верно оценивать их и т. п. Эти условия и их расшифровка – предмет других научных исследований. Философские размышления позволяют сделать вывод о том, что свобода и ответственность могут существовать в принципе. Он кажется интуитивно достоверным, но, может, тогда следовало сразу положиться на интуиции? Насколько важны размышления, приведенные в этой книге, если они приводят к выводу, совпадающему с исходными интуициями, если они соответствуют простому здравому смыслу?

Опора на интуиции важна как отправная точка размышлений. Но интуиций недостаточно в качестве основы для убеждений. Ведь несмотря на то, что интуиции часто достоверны и закреплены в языковой практике, иногда они все-таки оказываются ложными. Более того, как было видно из текста, многие ученые и некоторые философы оспаривают интуиции о свободе и ответственности. Они уверены, что доказали, почему свободы и ответственности быть не может. Размышления в этой книге опровергают аргументы критиков свободы и отстаивают ее возможность. В свою очередь, отрицание свободы не только безосновательно, но и вдобавок плохо влияет на мотивацию людей.

Убеждение в отсутствии свободы воли снижает желание

прилагать усилия и принимать сложные альтруистические решения. Для многих оно становится универсальным оправданием и поводом к деградации. Я могу подтвердить это, сославшись на несколько психологических исследований. Это эксперименты К. Вос и Дж. Скулера [Vohs, Schooler 2008]. Они показывают, что даже короткое знакомство с аргументами против свободы воли ослабляет моральные установки личности.

В одном таком эксперименте двум группам испытуемых предлагалось выполнить математические задачи. Во время выполнения заданий все участники могли сжульничать. До начала испытания участники первой группы читали параграфы из книги нобелевского лауреата Ф. Крика «Удивительная гипотеза: научный поиск души», в которых утверждалось, что большинство ученых считают убеждение в существовании свободы воли опровергнутым. Испытуемые в другой группе читали параграфы этой же книги, в которых ничего не говорится об отсутствии свободы воли. В результате эксперимента оказалось, что испытуемые, прочитавшие текст об отсутствии свободы воли, жульничали гораздо чаще.

В другом подобном эксперименте участникам предлагалось выполнять математические задания за вознаграждение. Вознаграждение полагалось за каждое правильно выполненное задание, однако их подсчет велся участниками самостоятельно. В результате этого эксперимента выяснилось, что больше правильно выполненных заданий засчитали себе те,

кто предварительно прочел текст об отсутствии свободы воли. Таким образом, Вос и Скулер показали, что убеждения в отсутствии свободы воли снижают моральные установки личности. Можно привести и другие аналогичные исследования. Большинство из них подтверждает негативное влияние убеждения об отсутствии свободы⁵². Следовательно, аргументы, изложенные выше, должны действовать наоборот, как противоядие, и мотивировать к моральному поведению. Кроме того, эти аргументы дают основание для эпистемологического оптимизма.

Иногда ученые и философы, которые не видят способа сочетать свободу, детерминизм, физикализм и событийный каузализм, начинают отрицать именно последние три тезиса. Наличие свободы представляется им уж слишком очевидным, и, по их мнению, они выбирают меньшее из двух зол. Однако это решение я считаю очень неудачным. Оно ведет к отходу от фундаментальных натуралистических, научных принципов. Но ведь именно на этих принципах основана большая часть знаний и достижений человечества. Детерминизм, физикализм и событийный каузализм позволяют считать, что мир однороден и события в нем совер-

⁵² Среди них можно отметить исследования Р. Баумейстера, свидетельствующие о том, что уменьшение веры в свободу воли влечет снижение альтруистического и рост агрессивного поведения [Baumeister, Masicampo, Dewall 2009]; работу Д. Ригони и его коллег, указывающую на снижение самоконтроля и рост склонности к антисоциальному поведению при снижении веры в свободу воли [Rigoni, Kühn et al. 2012] и др.

щаются законосообразно. Это, в свою очередь, дает основания считать его познаваемым в принципе. Отказавшись от этих тезисов, наука и философия рискуют заведомо признаться в своей несостоятельности и неспособности объяснить мир. Моя книга показывает, что оснований для такого отказа недостаточно, и направлена против *мистерианства*. Она демонстрирует, что свобода и ответственность не нуждаются в нематериальной субстанции, квантовой неопределенности и нарушении локальной супервентности и что не нужно искать лакуны в законах природы. Кроме того, книга проясняет ключевые философские понятия.

Прояснение терминов «причина», «свобода воли» и «ответственность» необходимо, поскольку от них непосредственно зависят выводы о возможности ментальной каузальности, свободы и ответственности. Бытовые значения этих слов слишком туманны, а философы предлагают различные, иногда не согласующиеся между собой определения. Из различных определений следуют различные, порой даже противоположные выводы. И утверждение о возможности ментальной каузальности, свободы и ответственности, к которому я прихожу в тексте, также обусловлено выработанными в ходе рассуждения формулировками. Почему я прихожу именно к этим определениям и нет ли в этом произвола?

Я использую метод последовательных итераций. Вначале я ориентируюсь на самые общие определения, практику словопотребления, и пытаюсь выявить заложенные в ней ин-

туиции. Эти интуиции в ходе анализа преобразуются в более точные философские формулировки, а затем сопоставляются с достоверными философскими утверждениями. Согласовав ключевые утверждения и объединив их в философскую позицию, я вновь уточняю понятия. Это позволяет мне утверждать предпочтительность принятых определений. Они ближе других к обыденному словоупотреблению, точнее схватывают исходные интуиции в понимании этих слов и согласуются с современным научным дискурсом. Из этого следует, что в моей книге не желаемая философская позиция определяет значения понятий, а исходные интуиции и практика словоупотребления. Полученный результат – это своеобразное *рефлексивное равновесие* между обыденным словоупотреблением, техническими определениями понятий и связанными с ними философскими убеждениями.

Итак, понятие причины – оно было одним из базовых кирпичиков при построении концепции ментальной каузальности. Первоначально я исходил из того, что высокоуровневые свойства, вводимые теориями специальных наук, обладают каузальной эффективностью. Думаю, с этим согласно большинство ученых и это соответствует позиции научного реализма. Высокоуровневые свойства обычно позволяют оптимальным образом объяснить регулярности событий. Ментальные состояния выполняют ту же роль и так же могут быть отнесены к разряду высокоуровневых свойств. Из этого следует возможность ментальной каузальности. Следстви-

ем этого умозаключения служит уточнение определения *каузальности как оптимального объяснения устойчивых регулярностей, следующих друг за другом и в непосредственной близости явлений*.

Конечно, это прагматистское, телеофункционалистское определение каузальности. Оно допускает каузальную эффективность высокоуровневых свойств, реализуемых в каких-то других низкоуровневых свойствах. Это определение согласуется с возможностью существования нескольких объяснений явления на разных уровнях, которые конкурируют между собой только в плане оптимальности или экономичности. И отрицает эксклюзивность физики как науки, в которой работают каузальные связи. Но это мне представляется лучшим определением причины, и к нему подводит логика рассуждений. Аналогичным путем я прихожу к определению свободы воли.

Изначально я предполагаю, что интуиции о свободе как о необходимом условии ответственности правильно схвачены в юридической практике. Согласно этой практике, ответственность возникает, когда агент действует осознанно, имея достаточно достоверной информации, без принуждения со стороны других агентов, и когда у него существуют альтернативы. Из этого следует, что *сущность свободы воли – это возможность совершать действия осмысленные, каузально обусловленные собственными желаниями агента при наличии у него возможности поступить иначе*. Как показыва-

ют рассуждения, представленные выше, наличие способности действовать сообразно желанию и возможности поступать иначе принципиально не противоречит ни детерминизму, ни индетерминизму. Ведь возможность поступать иначе предполагает, что агент мог поступить иначе в ситуациях *близких*, а не полностью идентичных.

Если свободу определить как наличие реальной возможности поступать иначе в точно такой же ситуации, свобода окажется несовместимой с детерминизмом, ведь детерминизм отрицает возможность разных последствий у абсолютно идентичных исходных событий. Но только это значение свободы – слишком узкое, и оно не имеет приложения к практическим дискуссиям. Именно поэтому стоит использовать определение, приведенное мной выше. Аналогично дистиллируется и понятие моральной ответственности. Это, пожалуй, самое важное и емкое в работе понятие. Его расшифровку я начинал с практики межличностного общения.

Обычно индикатором ответственности служат ответные реакции агентов. В отношении ответственного за поступок агента выносят моральные суждения, испытывают и выражают эмоциональные реакции, принимают встречные меры: наказывают или поощряют его. Но, как я заметил еще во Введении, в отношении других объектов эти реакции неуместны. Они нерациональны по отношению к силам природы и вызванным ими событиям: природные явления неуместно осуждать, на них глупо злиться, абсурдно пытаться их на-

казывать. Они неуместны в отношении неживых предметов и живых, но нерациональных существ. Поэтому описанные во Введении случаи не являются примерами ответственных действий. Из этого следует, что *ответственностью следует называть особого рода соотношение между действиями и агентом*. Это соотношение возникает при выполнении ряда условий. Во Введении я кратко перечислял их. В ходе работы они были расшифрованы, была произведена их редукция, а также получен ответ на вопрос, реализуемы ли они в принципе.

Одним из условий ответственности является *каузальная связь* между намерениями агента и его действиями. То есть желания и убеждения должны быть телеофункциональной причиной действий. Другим условием служит наличие *возможности* у агента поступать иначе, то есть возможности действовать различными способами в сходных обстоятельствах. Третьим условием служит возможность личности присваивать действия, включая их в свой нарратив. Это редуктивная база ответственности. И, как следует из размышлений, она не противоречит физикализму, детерминизму и событийному каузализму. Отсутствие такого противоречия говорит о возможности ответственности. Это стало одним из центральных выводов книги.

Однако этот вывод все еще не объясняет рациональности ответных реакций. Даже при наличии ответственности, этого специфического отношения между агентом и поступком,

ответные реакции могли бы быть нерациональными. Должны существовать дополнительные факторы, объясняющие, почему агенты реагируют на эти поступки соответствующим образом. Этими факторами выступают моральные представления агентов – ретрибутивистские или консеквенциалистские убеждения. Они объясняют веру агентов в то, что их реакции не только уместны, но значимы и справедливы. И теперь, в завершение книги, я хочу обратиться к анализу этих *рациональных оснований ответных реакций* в свете нарративной теории моральной ответственности, без чего представление о ней было бы не полным.

Ретрибутивизм – это убеждение, что наказание преступника – причинение ему боли, ограничение его прав и возможностей и даже в некоторых случаях его убийство – является благом, так как преступление само по себе заслуживает наказания. Заслуженность наказания в ретрибутивизме определяется справедливостью – некоторым моральным равновесием. В момент совершения преступления равновесие нарушается, возникает долг преступника перед обществом. Долг оплачивается через возмездие, если оно соразмерно совершенному аморальному или преступному действию. Тогда восстанавливается справедливость и равновесие. Ретрибутивистские соображения обосновывают осуждение и наказание за намеренный свободный аморальный поступок осведомленного преступника как восстановление справедливости, воздаяние по заслугам.

Иначе работает обоснование с помощью консеквенциализма. Согласно консеквенциалистским идеям (например, утилитаризму), осуждение и наказание преступника является благом в том случае, если они приводят к положительным последствиям (*consequences*). Положительным последствием является сокращение страданий в масштабах всего человечества: чем меньше будет страданий, тем лучше. Наказание само по себе благом не является, так как несет страдание преступнику. Но в инструментальном плане оно может быть благом, если в перспективе сокращает общую величину страданий. То есть в основе консеквенциалистской рационализации ответных реакций лежит не идея справедливости, а представление о пользе.

Любой тип ответных реакций, наказаний в том числе, может быть оправдан как ретрибутивистской, так и консеквенциалистской логикой. Моральные оценки и эмоциональные реакции – порицание, благодарность, гнев и обида – могут носить как ретрибутивное, так и утилитарное значение. Аналогичным образом дело обстоит с закрепленными в праве формами осуществления наказания: штрафами, ограничениями в правах, телесными наказаниями, тюремным заключением, исправительными работами и даже смертной казнью – все эти формы осуществления наказания в принципе согласуются как с ретрибутивизмом, так и консеквенциализмом. Например, тюремное заключение может быть формой воздаяния, то есть способом взыскать моральный долг

страданием за причиненные страдания другим. Но тюремное заключение может быть и способом обезопасить общество, оградить его от преступника и минимизировать общественные риски, то есть служить целям общего блага в духе утилитаризма. В этом отношении ретрибутивистская и утилитаристская логики зачастую подсказывают одинаковые практические решения. Однако в некоторых конкретных случаях различие в подходах приводит к предпочтению различных ответных реакций.

Если преступник полностью осознал свою вину, не имеет склонности или даже физической возможности совершать иные преступления и факт отсутствия наказания не повысит вероятность совершения других преступлений, исходя из консеквенциалистских соображений, его не следует наказывать. Ведь это не принесет никакой пользы, но только увеличит страдания. Однако, следуя ретрибутивистской логике, наказание за преступление в любом случае должно быть неминуемым, ведь вне зависимости от других последствий само восстановление справедливости и воздаяние являются благом. В некоторых случаях, по утилитаристской логике, допустимо несоответствующее преступлению наказание (более жестокое или более мягкое, вплоть до наказания невинных и отсутствия наказания виновных), если оно приведет к снижению преступлений в будущем. Согласно же ретрибутивистской логике, наказание всегда должно быть соразмерно тяжести и характеру преступления. Я считаю, что идеи, из-

ложенные в книге, лучше согласуются с консеквенциалистской позицией.

Нарративная концепция ответственности предполагает, что личность определяется автобиографической историей. Эта история состоит из распределенных во времени событий и их объяснений. Но разные главы текста этой истории не эквивалентны между собой. Личность как совокупность событий, состояний, желаний, убеждений, объединенных нарративом, не имеет постоянных характеристик, в том числе неизменной меры вины, морального долга. Каждая новая глава автобиографии дополняет и корректирует общую характеристику личности. Изменяются убеждения, желания, способность к волевому контролю, физические характеристики и культурный контекст. На основании лишь одной главы, в которой зафиксирован поступок, нерационально требовать воздаяния или выносить моральную оценку: новая глава автобиографии внесет корректировку в характеристики личности и значение ее поступка. Только в контексте всей истории, исходя из перспективы всей биографии, можно было бы определять размер «заслуги». Но история будет полной лишь после смерти личности, а значит, «текущий размер» заслуги личности не может быть зафиксирован окончательно. Если же его невозможно зафиксировать, единственным основанием для ответных реакций остается консеквенциализм.

Консеквенциалистские соображения в совокупности с от-

ветственностью работают как рациональные основания ответных реакций. Если благом является сокращение страданий, то рационально использовать ответные реакции в отношении ответственных действий. Только какие из этих ответных реакций наиболее эффективны? Какая система правосудия лучше соответствует утилитаристской логике? Какие меры будут наиболее эффективно предотвращать преступления и страдания? Эти практические вопросы я оставлю эпилогом к моей работе. У меня нет на них однозначного ответа. Но для того, чтобы определить возможные варианты, я приведу несколько сценариев. Это будут иллюстрации действия пенитенциарной системы из истории, современности и возможного будущего. Конечно, они далеко не исчерпывают всех возможных (удачных и неудачных) примеров практической реализации консеквенциалистской этики. Но они позволяют лучше понять ее сильные и слабые стороны и обозначить основные направления поиска наиболее удачного претворения ее в жизнь. Какая система, как вы думаете, наиболее эффективна?

Зрелищная казнь. Робер-Франсуа Дамьен был приговорен к казни четвертованием 28 марта 1757 г. В этот день его привезли на площадь Грев в мешке из замшевой кожи, в котором помещалось все тело преступника, кроме головы. На площади собралась толпа зрителей. Разодетые дамы предусмотрительно запасались нюхательной солью на случай обморока. В полчетвертого зрелище началось. Дамьена привя-

зали к платформе, а руку, совершившую преступление, жгли серой, пока Дамьен не издал ужасный вопль. Волосы у него встали дыбом. Эта часть казни продолжалась три минуты. Далее последовало вырывание клещами кусков мяса из тела преступника. В раны лили кипящее масло, горящую смолу, серу и расплавленный свинец. В воздухе витал запах жженой плоти. Потом Дамьена сняли с платформы и положили на деревянные брусья, сложенные в форме Андреевского креста.

Вокруг его предплечий, бедер, запястий и лодыжек обвязали крепкие веревки, в которые впрягли четырех сильных лошадей, по одной у каждого угла помоста. Взвился кнут, и лошади рванули вперед. Палачи ожидали, что лошади разорвут преступника на части. Но этого не произошло. Целый час лошадей стегали и пришпоривали, но им никак не удавалось оторвать хотя бы одну ногу или руку. Дамьен душераздирающе кричал. К нему несколько раз с утешениями подходили священники. После множества неудачных попыток четвертования палачи получили от судей разрешение перерезать преступнику связки и раздробить суставы, чтобы облегчить работу лошадям. Затем на животных вновь обрушились удары кнута, и через полтора часа от начала казни они, наконец, смогли оторвать сначала правую, а потом и левую ногу Дамьена. Люди на площади захлопали. После того как все конечности преступника были отделены от туловища, палач сбрил с головы Дамьена волосы. Их вместе с оторванными членами бросили в костер, где потом сожгли и туловище

еще живого, как говорят, преступника. Пепел Дамьена был развеян по ветру.

Скрытый изолятор на всю жизнь. Заключенный камеры № 174 приговорен к пожизненному сроку в 1997 г. Он никогда не видел собственной тюрьмы со стороны. Арестантов привозят сюда с мешками на головах и заводят под лай собак внутрь. В коридорах тюремного корпуса – стерильная чистота и тишина. В камере установлены прикрученные к полу двухъярусные кровати, стол, скамейка, бак с питьевой водой, вешалка для одежды, полка, рукомойник и унитаз. Камера длиной в пять с половиной шагов. Свет здесь никогда не выключается, как и видеонаблюдение. В камере вместе сидят по четыре человека. Спят они головой к двери, не накрывая лица, руки поверх одеяла. Решетки отделяют заключенных от наружной двери и окна.

Как только в камеру открывается форточка или дверь, заключенные должны принять позу «ку»: встать к стене, широко расставить ноги, низко наклониться, оперевшись в стену затылком, руки поднять за спиной вверх, словно размахиваясь для прыжка в воду, пальцы растопырить, глаза закрыть, а рот, наоборот, широко открыть. Таким образом, надзиратели видят, что никаких запрещенных предметов нет. Проверки происходят по нескольку раз в день. Во время проверок дежурный по камере рапортует: имена заключенных, уголовные статьи и коротко – за какие преступления отбывают срок. Передвижение даже между двумя ближайшими

помещениями допустимо лишь в присутствии двух охранников и кинолога с собакой. Оно происходит в той же самой позе – позе «ку». Собака натренирована лаять, пока заключенный находится вне камеры. Любое нарушение – плохо заправленная койка или недостаточно широко раскрытый рот – расценивается как нападение на охрану. Такие вольности караются газом «Черемуха» и резиновой дубинкой.

Каждый день выполняется абсолютно одинаковый распорядок. В 6 утра – подъем. Заключенные аккуратно заправляют кровати, умываются, проводят уборку в камере. В 6:30 – завтрак. На завтрак – сечка, перловая, ячневая или овсяная каша. С 7:15 до 8:00 работают электророзетки. Можно побриться или вскипятить воду. С 6:15 до 8:00 работает радио. В 8:00 – утренняя проверка и вывод на работу. Осужденные могут заниматься простым физическим трудом. С 10:00 – прогулка в каменном дворике. С 13:00 до 13:50 – обед. Затем – десять минут на уборку камер. После обеда – продолжение работы. С 14:00 до 17:00 вновь включается радио. В 17:00 – съем с работы и вывод второй смены. С 18:00 до 18:40 – ужин. С 20:00 до 20:30 – вечерняя проверка. С 20:30 вновь включают электророзетки и радио. Выключают незадолго перед отбоем. В 22:00 – отбой для первой смены. Съем с работы второй смены осуществляется в 3:00 утра. И так – всю оставшуюся жизнь.

Экоотель, где можно учиться и работать. Некоторые особо опасные преступники из Норвегии попадают на ост-

ров Бастой. Там находится самая райская из ныне существующих в мире тюрем. Расположенный всего в часе езды от Осло остров представляет собой живописное место с пляжами, теннисными кортами, сауной и мастерскими. заключенные этой тюрьмы проживают в уютных деревянных домиках на берегу моря по 8 человек. У каждого своя комната. Они без ограничений передвигаются по острову, играют в футбол, хоккей, загорают. Один из них так отзывается о своем пребывании на острове Бастой: «Я не могу сказать, что мне нравится находиться здесь. Но если бы это не было тюрьмой, норвежское правительство могло бы неплохо зарабатывать на этом комплексе, сдавая номера в качестве гостиничных». Удивительно, но это одна из самых дешевых тюрем для государственного бюджета. 120 заключенных охраняют и обслуживают всего семьдесят сотрудников, все без оружия. Только пять или шесть из них ночуют на острове.

Такой экономии правительство добилось за счет того, что практически все работы по поддержанию колонии выполняют сами заключенные. Они убирают мусор, готовят, рубят лес, изготавливают инструменты и мебель, выращивают продукты и ловят рыбу. Именно труд и обучение являются главными обязанностями заключенных. На работе они находятся с 8:30 до 15:30. За это они получают жалованье в размере 10 долларов в день, которое могут потратить на продукты в местном магазине. Кроме жалованья, за труд им положен отпуск – 18 дней в году, и еще... декретный, если у жены

или сожительницы рождается ребенок. Такое, кстати, случается достаточно часто, ведь свидания с родственниками у них еженедельные, по 12 часов в неделю.

*Дивный новый мир*⁵³. В 2054 г. было создано отделение предотвращения преступлений. Благодаря работе отделения удалось свести число фактических убийств до нуля. Это радикально изменило систему правосудия: исчезли виновные в тяжких преступлениях, исчезли жертвы преступлений, но появились предпреступники – лица, которые должны были совершить преступления. Работа по предотвращению преступлений основывается на компьютерной системе мониторинга и прогнозирования, которая за несколько дней до преступления с абсолютной точностью предсказывает, что в отсутствие вмешательства оно неизбежно произойдет. Программа идентифицирует жертву и ее будущего убийцу. После чего сотрудники отделения задерживают предпреступника и помещают в изолятор, где его подключают к виртуальной реальности. В изоляторе поддерживается жизнь тела преступника, а он сам существует в мире симуляции до самой старости. В этой виртуальной реальности он изолирован от реального взаимодействия с другими людьми и не может причинить кому-либо зла, при этом сам он этого не осознает. Он общается с виртуальными людьми, строит свою жизнь, карьеру, а иногда совершает преступления, за которые его

⁵³ Эта вымышленная альтернатива основана на рассказе американского писателя-фантаста Филипа Дика «Особое мнение».

не наказывают. С точки зрения создателей и сторонников такой системы предотвращения преступлений, она наиболее эффективна в плане минимизации страданий: как жертв, так и самих преступников.

Рига, 02.04.2018

Библиография

Тексты на русском языке

1. *Беседин А. П.* 2015. Натуралистический аргумент П. Стросона в пользу моральной ответственности // Вестник Московского университета. Серия 7: Философия. № 6. С. 8–16.
2. *Беседин А. П.* 2016. Может ли детерминизм нами манипулировать // Логос. № 5. С. 201–212.
3. *Беседин А. П., Васильев В. В., Волков Д. Б., Кузнецов А. В.* 2017. О чем думают российские философы? Результаты интернет-опроса // Вестник Московского университета. Серия 7: Философия. № 6. С. 71–83.
4. *Брентано Ф.* 1996. Психология с эмпирической точки зрения // Ф. Брентано. Избранные работы. М.: Дом интеллектуальной книги, Русское феноменологическое общество. С. 9–91.
5. *Васильев В. В.* 2009. Трудная проблема сознания. М.: Прогресс-Традиция.
6. *Васильев В. В.* 2014. Сознание и вещи: Очерк феноменалистической онтологии. М.: Книжный дом Либроком.
7. *Васильев В. В.* 2016. В защиту классического компатибилизма // Вопросы философии. № 2. С. 64–76.

8. *Васильев В. В.* 2017. В защиту классического компатибилизма: Эссе о свободе воли. М.: Ленанд.
9. *Волков Д. Б.* 2007. Споры о сознании и аналитическая философия // Эпистемология и философия науки. № 3. С. 176–181.
10. *Волков Д. Б.* 2012. Бостонский зомби: Д. Деннет и его теория сознания. М.: Либроком.
11. *Волков Д. Б.* 2015а. Опровергает ли Аргумент каузальных траекторий локальную супервентность ментального над физическим // Эпистемология и философия науки. Т. 2 (44). С. 166–182.
12. *Волков Д. Б.* 2015б. Свобода воли в либертарианстве Кейна // Вестник Московского университета. Серия 7: Философия. № 6. С. 32–51.
13. *Волков Д. Б.* 2015с. Что доказывают манипуляции с «Аргументом манипуляции» Д. Перебума? // Философия и культура. № 6. С. 933–942.
14. *Волков Д. Б.* 2016а. Являются ли психопаты морально ответственными личностями? // Вестник Томского государственного университета. Философия. Социология. Политология. № 2 (34). С. 212–222.
15. *Волков Д. Б.* 2016б. Проблема свободы воли. Обзор ключевых исследований конца XX – начала XXI вв. в аналитической философии // Философский журнал. Т. 9. № 3. С. 175–189.
16. *Волков Д. Б.* 2016с. Нарративный подход как решение

вопроса о тождестве личности // Вестник СПбГУ. Сер. 17. Философия. Конфликтология. Культурология. Религиоведение. № 4. С. 21–32.

17. Волков Д. Б. 2016d. Рецензия на книгу Д. Перебума «Свобода воли, агент и смысл жизни» // Вестник Томского государственного университета. Философия. Социология. Политология. № 4 (36). С. 404–409.

18. Волков Д. Б. 2016e. «Тезис о сознании» и моральная ответственность в исследованиях Нила Леви // Логос. № 5. С. 213–226.

19. Волков Д. Б. 2016f. Аргумент исключения Дж. Кима и проблема ментальной каузальности // Вестник Московского университета. Серия 7: Философия. № 6. С. 15–33.

20. Волков Д. Б. 2016g. Проблема ментальной каузальности: обзор новейших исследований // Философия и культура. № 6. С. 673–682.

21. Волков Д. Б. 2017a. Современный субстанциональный подход к проблеме тождества личности // Философия и культура. № 1. С. 77–85.

22. Волков Д. Б. 2017b. Решение проблемы ментальной каузальности в биологическом натурализме Дж. Сёрла // Философская мысль. № 2. С. 1–12.

23. Волков Д. Б. 2017c. Психологический подход как решение проблемы тождества личности // Философская мысль. № 2. С. 13–25.

24. Волков Д. Б. 2018. Преимущества нарративного под-

хода к проблеме тождества личности // *Философский журнал*. Т. 11. № 3. С. 166–175.

25. *Волков Д. Б., Деннет Д.* 2015. Об организации конференции в Гренландии «Проблемы сознания и свободы воли в аналитической философии» // *Вестник Московского университета*. Серия 7: Философия. № 6. С. 3–7.

26. *Гаспаров И. Г.* 2018. Тождество личности и нарратив // *Философский журнал*. Т. 11. № 3. С. 180–183.

27. *Дубровский Д. И.* 1983. Проблема идеального. Субъективная реальность. М.: Мысль.

28. *Дубровский Д. И.* 2007. Сознание, мозг, искусственный интеллект. М.: ИД Стратегия-Центр.

29. *Иванов Д. В.* 2013. Природа феноменального сознания. М.: Либроком.

30. *Кузнецов А. В.* 2015. Конститутивный панпсихизм и проблема ментальной каузальности // *Вестник Московского университета*. Серия 7: Философия. № 6. С. 106–117.

31. *Кузнецов А. В.* 2016а. Локальный интеракционизм. Новое решение проблемы ментальной каузальности // *Вопросы философии*. № 1. С. 148–161.

32. *Кузнецов А. В.* 2016б. Проблема ментальной каузальности в аналитической философии сознания. Диссертация на соискание степени кандидата философских наук: 09.00.03 – История философии. Москва.

33. *Левин С. М.* 2013. Физикализм, субъект и метафизическая свобода воли // *Вестник Ленинградского государственного университета имени Александра Ф. Керенского*. Т. 10. № 1. С. 10–17.

ного университета имени А. С. Пушкина. № 2. С. 67–77.

34. *Левин С. М.* 2016а. Коллективное сознание и моральная ответственность // Вестник Томского государственного университета. Философия. Социология. Политология. № 3. С. 46–53.

35. *Левин С. М.* 2016b. Моральная ответственность и личность корпораций: нарративный подход // Мысль: Журнал Петербургского философского общества. № 21. С. 63–76.

36. *Левин С. М.* 2018. Нарративный подход: фикция и реальность // Философский журнал. Т. 11. № 3. С. 184–187.

37. *Лейбниц Г. В.* 1982. Сочинения в четырех томах. Т. 1. М.: Мысль.

38. *Локк Дж.* 1985. Сочинения в 3-х т. Т. 1. Опыт о человеческом разумении. М.: Мысль.

39. *Мареева Е. В.* 2003. Дж. Сёрл: старое и новое в понятии сознания // Философия сознания: история и современность. С. 216–231.

40. *Молчанов В. И.* 2007. Исследования по феноменологии сознания. М.: Издательский дом «Территория будущего».

41. *Нагуманова С. Ф.* 2011. Материализм и сознание: анализ дискуссии о природе сознания в современной аналитической философии. Казань: Казан. ун-т.

42. Проблема сознания в философии и науке. 2009 / Под ред. Д. И. Дубровского. М.: Канон+.

43. *Разеев Д. Н.* 2015. Ложные воспоминания и тождество

личности // Вопросы философии. № 1. С. 45–50.

44. *Разин А. В.* 2015. Свобода и необходимость // Вестник Российского университета дружбы народов. Серия: Философия. № 2. С. 9–15.

45. *Секацкая М. А.* 2018. Нарративный подход и проблема двойников // Философский журнал. Т. 11. № 3. С. 176–179.

46. *Секацкая М. А.* 2010. Физический детерминизм и свобода воли: картезианское решение // Философские науки. № 10. С. 65–80.

47. *Секацкая М. А.* 2012. Этические идеалы, логические ограничения и проблема свободы // Вопросы философии. № 2. С. 81–91.

48. *Секацкая М. А.* 2014. Пересадка мозга и тождество личности: альтернативная интерпретация одного мысленного эксперимента // Эпистемология и философия науки. Т. 42. № 4. С. 67–76.

49. *Секацкая М. А.* 2016а. Свобода воли и предсказуемость. Философский анализ современных исследований в нейронауке // Вопросы философии. № 3. С. 163–169.

50. *Секацкая М. А.* 2016б. Интенциональность, свобода воли и моральная ответственность коллективов // Журнал Петербургского философского общества. № 21. С. 77–86.

51. *Сёрл Дж.* 2002. Открывая сознание заново. М.: Идея-Пресс.

52. *Сёрль Дж.* 2004. Рациональность в действии. М.: Прогресс-Традиция.

53. *Скворцов А. А.* 2016. *Этика: учебник для бакалавров / Под общ. ред. А. А. Гусейнова.* М.: Издательство Юрайт.
54. *Уголовное право России. 2004. Части Общая и Особенная: учеб. / Под ред. А. И. Рарога.* М.: ИНФРА-М.
55. *Хофштадтер Д., Деннетт Д.* 2003. *Глаз разума.* Самара: Издательский Дом «Бахрах-М».
56. *Чирва Д. В.* 2015. «Правильная» постановка вопроса о свободе: деятельностный индетерминизм // *Этносоциум и межнациональная культура.* № 10 (88). С. 87–93.
57. *Этика. 2016. Учебник для академического бакалавриата / Под общ. ред. А. А. Гусейнова.* М.: Издательство Юрайт.
58. *Юлина Н. С.* 2007. *Философский натурализм: о книге Дэниела Деннета «Свобода эволюционирует».* М.: Канон+.
59. *Юм Д.* 1996а. *Сочинения в 2 т. Т. 1.* М.: Мысль.
60. *Юм Д.* 1996б. *Сочинения в 2 т. Т. 2.* М.: Мысль.
61. *Юм Д.* 2009. *Исследование о человеческом познании.* М.: «Канон+» РООИ «Реабилитация».

Тексты на английской языке

1. *Akelaitis A.* 1945. Studies on the Corpus Callosum: IV Diagnostic Dyspraxia in Epileptics Following Partial and Complete Section of the Corpus Callosum // The American Journal of Psychiatry. Vol. 101. Issue 5. P. 594–599.
2. *Atkins K.* 2010. Narrative Identity and Moral Identity. A Practical Perspective. Oxford: Routledge.
3. *Balaguer M.* 2010. Free Will as an Open Scientific Problem. Cambridge: A Bradford Book.
4. *Baumeister R.F., Masicampo E.J., Dewall C.N.* 2009. Prosocial Benefits of Feeling Free: Disbelief in Free Will Increases Aggression and Reduces Helpfulness // Personality & Social Psychology Bulletin. Vol. 35. Issue 2. P. 260–268.
5. *Block N.* 1990. Can the Mind Change the World? // Boolos G. (ed.). Meaning and Method: Essays in Honor of Hilary Putnam. Cambridge: Cambridge University Press. P. 137–170.
6. *Block N.* 2003. Do Causal Powers Drain Away? // Philosophy and Phenomenological Research. Vol. 67. Issue 1. P. 133–150.
7. *Brentano F.* 1924. Psychologie vom empirischen Standpunkt. Erster Band. Leipzig: F. Meiner.
8. *Chalmers D.* 1997. The Conscious Mind. In Search of a Fundamental Theory. Oxford, New York: Oxford University Press.

9. *Chisholm R.* 1964. Human Freedom and The Self. The Lindley Lecture, University of Kansas.
10. *Chisholm R.* 1991. On the Simplicity of the Soul // Philosophical Perspectives, Vol. 5, Philosophy of Religion. P. 167–181.
11. *Christman J.* 2004. Narrative Unity as a Condition of Personhood // *Metaphilosophy*. Vol. 35. № 5. P. 695–713.
12. *Churchland P.* 2013. Touching a Nerve: Our Brains, Our Selves. New York.
13. *Colyvan M.* 1998. Can the Eleatic Principle Be Justified? // *Canadian Journal of Philosophy*. Vol. 28. № 3. P. 313–335.
14. *Cornell D., Warren J., Hawk G.* et al. 1996. Psychopathy in Instrumental and Reactive Violent Offenders // *Journal of Consulting and Clinical Psychology*. Vol. 64. № 4. P. 783–790.
15. *Crane T.* 1992. Mental Causation and Mental Reality // *Proceedings of the Aristotelian Society*. Vol. 92. P. 185–202.
16. *Crane T.* 2001. Elements of Mind: An Introduction to the Philosophy of Mind. Oxford: Oxford University Press.
17. *Davidson D.* 1980. Essays on Actions and Events. Oxford: Clarendon Press.
18. *Davidson D.* 2006a. Action, Reasons, and Causes // *The Essential Davidson*. Oxford, New York: Oxford University Press. P. 23–36.
19. *Davidson D.* 2006b. Mental Events // *The Essential Davidson*. Oxford, New York: Oxford University Press. P. 105–

121.

20. *Dennett D.* 1976. *Conditions of Personhood* // Rorty A (ed.). *The Identities of Persons*. Berkley: University of California Press. P. 175–196.

21. *Dennett D.* 1981. *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge, Massachusetts: The MIT Press.

22. *Dennett D.* 1984. *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, Mass.: MIT Press.

23. *Dennett D.* 1991a. *Real Patterns* // *The Journal of Philosophy*. Vol. 88. № 1. P. 27–51.

24. *Dennett D.* 1991b. *Consciousness Explained*. Boston, Massachusetts: Little Brown & Co.

25. *Dennett D.* 1992. *The Self as a Center of Narrative Gravity* // Kessel F., Cole P. and Johnson D. (eds.) *Self and Consciousness: Multiple Perspectives*. Hillsdale, NJ: Erlbaum. P. 103–115.

26. *Dennett D.* 1995. *Darwin's Dangerous Ideas. Evolution and the Meaning of Life*. New York: Touchstone.

27. *Dennett D.* 1996. *The Intentional Stance*. Cambridge, Massachusetts: The MIT Press.

28. *Dennett D.* 1998. *Brainchildren: Essays on Designing Minds*. Cambridge, Massachusetts: The MIT Press.

29. *Dennett D.* 2003. *Freedom Evolves*. N.Y.: Penguin Books.

30. *Dennett D.* 2010. *Two Black Boxes: A Fable* // *Activitas Nervosa Superior*. Vol. 52. № 2. P. 81–84.

31. *Dennett D.* 2013. *Intuition Pumps and Other Tools for Thinking*. N.Y.: W.W. Norton & Company.
32. *Dennett D.* 2017. *From Bacteria to Back and Back. The Evolution of Minds*. N.Y.: Norton and Company.
33. *Dennett D., Taylor C.* 2011. *Who's Still Afraid of Determinism? Rethinking Causes and Possibilities* // Kane R. (ed.). *The Oxford Handbook of Free Will: Second Edition* (2 ed.). Oxford: Oxford University Press. P. 221–242.
34. *Dretske F.* 1991. *Explaining Behavior: Reasons in a World of Causes*. Cambridge, MA: MIT Press.
35. *Dretske F.* 1993. *Mental Events as Structuring Causes of Behavior* // Heil J. and Mele A. (eds.). *Mental Causation*. Oxford: Clarendon Press. P. 121–136.
36. *Duff A.* 1977. *Psychopathy and Moral Understanding* // *American Philosophical Quarterly*. Vol. 14. № 3. P. 189–200.
37. *Earman J.* 1986. *A Primer on Determinism*. Boston: Springer.
38. *Ehring D.* 2009. *Causal Relata* // H. Beebe, C. Hitchcock, and P. Menzies (eds.). *The Oxford Handbook of Causation*. Oxford, New York: Oxford University Press. P. 387–413.
39. *Ekstrom L.* 2010. *Volition and the Will* // T. O'Connor and C. Sandis (eds.). *A Companion to the Philosophy of Action*. New Jersey: Wiley Blackwell. P. 99–107.
40. *Ekstrom L.* 2011. *Free Will is Not a Mystery* // Kane R. (ed.). *The Oxford Handbook of Free Will: Second Edition* (2 ed.). Oxford: Oxford University Press. P. 366–380.

41. *Fischer J.* 1986. *Moral Responsibility*. Ithaca: Cornell University Press.
42. *Fischer J.* 1994. *The Metaphysics of Free Will: An Essay on Control*. Cambridge, MA: Wiley-Blackwell.
43. *Fischer J.* 1999. Recent Work on Moral Responsibility // *Ethics*. № 110 (10). P. 93–139.
44. *Fischer J.* 2006. *My Way: Essays on Moral Responsibility*. Oxford: Oxford University Press.
45. *Fischer J.* 2010. The Frankfurt Cases: The Moral of the Stories // *Philosophical Review*. Vol. 119. № 3. P. 315–336.
46. *Fischer J.* 2011. Frankfurt-Type Examples and Semi-Compatibilism // Kane R. (ed.), *The Oxford Handbook of Free Will: Second Edition (2 ed.)*. Oxford: Oxford University Press. P. 243–265.
47. *Fischer J.* 2012. *Deep Control: Essays on Free Will and Value*. Oxford: Oxford University Press.
48. *Fischer J., Ravizza M.* 1993. *Perspectives on Moral Responsibility*. Ithaca: Cornell University Press.
49. *Fischer J., Ravizza M.* 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
50. *Fodor J.* 1981. The Mind-Body Problem // *Scientific American*. № 244. P. 114–125.
51. *Frankfurt H.* 1988. *The Importance of What We Care About: Philosophical Essays*. Cambridge: Cambridge University Press.

52. *Frankfurt H.* 2013. Alternate Possibilities and Moral Responsibility // P. Russell and O. Deery (eds.). *The Philosophy of Free Will: Essential Readings From the Contemporary Debates*. Oxford, New York: Oxford University Press. P. 139–148.

53. *Franklin C.* 2014. Event-Causal Libertarianism, Functional Reduction and the Disappearing Agent Argument // *Philosophical Studies*. Vol. 170. Issue 3. P. 413–432.

54. *Gazzaniga M.* 2011. *Who's in Charge? Free Will and the Science of the Brain*. New York: Ecco.

55. *Ginet C.* 1990. *On Action*. Cambridge: Cambridge University Press.

56. *Ginet C.* 2002. Reasons Explanations of Action: Causalist Versus Noncausalist Accounts // Kane R. (ed.) *The Oxford Handbook of Free Will*. Oxford University Press. P. 386–387.

57. *Goodman M.* 1992. A Sufficient Condition for Personhood // *The Personalist Forum*. Vol. 8 (Supplement). No 1. P. 75–81.

58. *Haji I.* 2001. Control Conundrums: Modest Libertarianism, Responsibility and Explanation // *Pacific Philosophical Quarterly*. Vol. 82. No 2. P. 178–200.

59. *Hardcastle V.G.* 1998. On the Matter of Minds and Mental Causation // *Philosophy and Phenomenological Research*. No 58. P. 1–25.

60. *Hare R.* 1993. *Without Conscience: The Disturbing World of the Psychopaths Among Us*. NY, London: The Guilford Press.

61. *Hart H.L.A.* 2008. *Punishment and Responsibility: Essays in the Philosophy of Law.* Oxford.
62. *Jackson F.* 1995. Essentialism, Mental Properties and Causation // *Proceedings of the Aristotelian Society.* Vol. 69. P. 253–268.
63. *Jackson F.* 1996. Mental Causation // *Mind.* № 105. P. 377–413.
64. *James R., Blair R.* 2003. Neurobiological Basis of Psychopathy // *The British Journal of Psychiatry.* № 182. P. 5–7.
65. *Kane R.* 1996. *The Significance of Free Will.* Oxford, N.Y.: Oxford University Press.
66. *Kane R.* 2007. *Libertarianism* // *Fischer J.M., Kane R., Pereboom D., Vargas M.* *Four Views on Free Will.* New Jersey: Willey Blackwell. P. 5–43.
67. *Kane R.* 2011a. Introduction: The Contours of Contemporary Free-Will Debates (Part 2) // *Kane R. (ed.). The Oxford Handbook of Free Will: Second Edition (2 ed.).* Oxford: Oxford University Press. P. 3–37.
68. *Kane R.* 2011b. Rethinking Free Will: New Perspectives on an Ancient Problem // *Kane R. (ed.). The Oxford Handbook of Free Will: Second Edition (2 ed.).* Oxford: Oxford University Press. P. 381–404.
69. *Kane R.* 2013. Responsibility, Luck and Chance: Reflections on Free Will and Indeterminism // *Russell P. and Deery O. (eds.). The Philosophy of Free Will: Essential Readings From the Contemporary Debates.* Oxford, New York: Oxford

University Press. P. 193–214.

70. *Kim J.* 1993a. *Supervenience and Mind: Selected Philosophical Essays.* Cambridge: Cambridge University Press.

71. *Kim J.* 1993b. The Non-Reductivist's Troubles with Mental Causation // Heil J. and Mele A. (eds.). *Mental Causation.* Oxford: Clarendon Press. P. 189–210.

72. *Kim J.* 2000. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation.* Cambridge MA: Bradford Book.

73. *Kim J.* 2006. *Philosophy of Mind.* 2nd edn. Oxford, Boulder: Westview Press.

74. *Kim J.* 2008. *Physicalism, or Something Near Enough.* Princeton: Princeton University Press.

75. *Kim J.* 2010. *Essays in the Metaphysics of Mind.* Oxford: Oxford University Press.

76. *Kind A.* 2015. *Persons and Personal Identity.* Malden: Polity Press.

77. *Klausen E., Passman R.* 2006. Pretend Companions (Imaginary Playmates): The Emergence of a Field // *The Journal of Genetic Psychology: Research and Theory on Human Development.* Vol. 167. Issue 4. P. 349–364.

78. *Lepore E., Loewer B.* 1987. *Mind Matters* // *Journal of Philosophy.* № 84. P. 630–642.

79. *Lepore E., Loewer B.* 1989. More on Making Mind Matter // *Philosophical Topics.* № 18. P. 175–191.

80. *Lepore E., Ludwig K.* 2006. Introduction // *The Essential*

Davidson. Oxford: Oxford University Press. P. 1–20.

81. *Levine J.* 2001. *Purple Haze: The Puzzle of Consciousness.* New York: Oxford University Press.

82. *Levy N.* 2005. *The Good, the Bad, and the Blameworthy // Journal of Ethics and Social Philosophy.* Vol. 1, № 2. P. 2–16.

83. *Levy N.* 2007. *The Responsibility of the Psychopath Revisited // Philosophy, Psychiatry, and Psychology.* Vol. 14. P. 129–138.

84. *Libet B.* 1985. *Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action // Behavioral and Brain Sciences.* № 8. P. 529–566.

85. *Libet B.* 2001. *Consciousness, Free Action and the Brain // Journal of Consciousness Studies.* № 8. P. 59–65.

86. *Locke J.* 2008. *Of Identity and Diversity // Perry J. (ed.). Personal Identity.* Berkeley: University of California Press. P. 31–52.

87. *Lowe E.* 1992. *The Problem of Psychophysical Causation // Australasian Journal of Philosophy.* № 70. P. 263–276.

88. *MacIntyre A.* 2007. *After virtue: A Study in Moral Theory.* Notre Dame, Indiana: University of Notre Dame Press.

89. *Malcolm N.* 1968. *The Conceivability of Mechanism // Philosophical Review.* № 77. P. 45–72.

90. *McKenna M.* 1998. *The Limits of Evil and the Role of Moral Address: A Defense of Strawsonian Compatibilism // Journal of Ethics.* № 2. P. 123–142.

91. *McKenna M.* 2008. A Hard-Line Reply to Pereboom's Four-Case Manipulation Argument // *Philosophy and Phenomenological Research*. Vol. 77. № 1. P. 142–159.
92. *McKenna M.* 2012. *Conversation and Responsibility*. N.Y.: Oxford University Press.
93. *McKenna M., Pereboom D.* 2016. *Free will: A Contemporary Introduction*. N.Y.: Routledge Taylor and Francis Group.
94. *McKenna M., Russell P.* 2008. *Free Will and Reactive Attitudes: Perspectives on P.F. Strawson's «Freedom and Resentment»*. Burlington, VT: Ashgate Publishing.
95. *Mele A.* 2005. A Critique of Pereboom's 'Four-Case Argument' for Incompatibilism // *Analysis*. Vol. 65. № 285. P. 75–80.
96. *Mele A.* 2008. *Free Will and Luck*. New York: Oxford University Press.
97. *Mele A.* 2009. Causation, Action, and Free Will // H. Beebe, C. Hitchcock, and P. Menzies (eds.). *The Oxford Handbook of Causation*. Oxford, New York: Oxford University Press. P. 554–574.
98. *Mele A.* 2013. *Dialog on Free Will and Science*. Oxford University Press.
99. *Mele A.* 2017. On Pereboom's Disappearing Agent Argument // *Criminal Law and Philosophy*. Vol. 11. № 3. P. 561–574.
100. *Misak C.* 2005. *Verificationism. Its History and*

Prospects. London and New York: Routledge.

101. *Murphy J.* 1972. Moral Death: A Kantian Essay on Psychopathy // *Ethics*. Vol. 82. № 4. P. 284–298.

102. *Nagel T.* 1986. *The View From Nowhere*. New York: Oxford University Press.

103. *Nelkin D.* 2007. Do We Have a Coherent Set of Intuitions about Moral Responsibility? // *Midwest Studies in Philosophy*. № 31. P. 243–259.

104. *Nelkin D.* 2011. *Making Sense of Freedom and Responsibility*. New York: Oxford University Press.

105. *Noonan H.* 2003. *Personal Identity*. London and New York: Routledge.

106. *Nozick R.* 1983. *Philosophical Explanations*. Cambridge, MA: Harvard University Press.

107. *Olson E.* 1997. *The Human Animal. Personal Identity Without Psychology*. Oxford: Oxford University Press.

108. *Olson E.* 2017. *Personal Identity* // *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition) / Edward N. Zalta (ed.). URL: <<https://plato.stanford.edu/archives/sum2017/entries/identity-personal/>>.

109. *Otten J.* 1977. Reviving the Logical Connection Argument // *Canadian Journal of Philosophy*. Vol. 7. № 4. P. 725–743.

110. *Parfit D.* 1971. *Personal Identity* // *The Philosophical Review*. Vol. 80, № 1. P. 3–27.

111. *Parfit D.* 1986. *Reasons and Persons*. Oxford: Clarendon

Press.

112. *Pereboom D.* 2001. *Living Without Free Will*. Cambridge: Cambridge University Press.

113. *Pereboom D.* 2008. *A Hard-Line Reply to the Multiple-Case Manipulation Argument // Philosophy and Phenomenological Research*. Vol. 77. № 1. P. 160–170.

114. *Pereboom D.* 2013a. *Optimistic Skepticism about Free Will // Russell P. and Deery O. (eds.). The Philosophy of Free Will: Essential Readings From the Contemporary Debates*. Oxford, New York: Oxford University Press. P. 421–449.

115. *Pereboom D.* 2013b. *Free Will // Crisp R. (ed.). The Oxford Handbook of the History of Ethics*. Oxford: Oxford University Press. P. 606–637.

116. *Pereboom D.* 2014a. *Responses to John Martin Fischer and Dana Nelkin // Science, Religion, and Culture*. Vol. 1. № 3. P. 218–225.

117. *Pereboom D.* 2014b. *Free Will, Agency, and Meaning in Life*. N.Y: Oxford University press.

118. *Pereboom D.* 2017. *Replies to Victor Tadros, Saul Smilansky, Michael McKenna, and Al Mele // Criminal Law and Philosophy*. Vol. 11. № 3. P. 617–636.

119. *Perry J.* 1978. *A Dialogue on Personal Identity and Immortality*. Indianapolis: Hackett Publishing.

120. *Perry J.* 2002. *Identity, Personal Identity, and the Self*. Indianapolis, Cambridge: Hackett Publishing Company.

121. *Pinker S.* 2002. *The Bank Slate: The Modern Denial of*

Human Nature. London: Penguin Books.

122. *Psillos S.* 2009. Regularity Theories // Beebe H., Hitchcock C., and Menzies P. (eds.). Oxford Handbook of Causation. Oxford, New York: Oxford University Press. P. 131–157.

123. *Putnam H.* 1967. Psychological Predicates // Capitan W.H. and Merrill D.D. (eds.). Art, Mind, and Religion. University of Pittsburgh Press. P. 37–48.

124. *Reid M.* 1997. Narrative and Fission: A Review of Marya Schechtman's *The Constitution of Selves* // Philosophical Psychology. Vol. 10. № 2. P. 211–219.

125. *Reid T.* 1969. Essay on the Intellectual Power of Man. Boston: MIT Press.

126. *Reid T.* 2008. Of Mr. Locke's Account of our Personal Identity // Perry J. (ed.). Personal Identity. Berkeley: University of California Press. P. 113–118.

127. *Rigoni D., Kühn S., et al.* 2012. Reducing Self-Control by Weakening Belief in Free Will // Consciousness and Cognition. Vol. 21. P. 1482–1490.

128. *Roberts J.* 1989. Aristotle on Responsibility for Action and Character // Ancient Philosophy. Vol. 9. № 1. P. 23–36.

129. *Ross D.* 2000. Rainforest Realism: A Dennettian Theory of Existence // Ross D., Brook A., Thompson D. (eds.) *Dennett's Philosophy: A Comprehensive Assessment*. Cambridge: MIT Press. P. 147–166.

130. *Sapolsky R.* 2004. The Frontal Cortex and the Criminal

Justice System // Philosophical Transactions of the Royal Society of London B. Biological Science. Vol. 359. No 1451. P. 1787–1796.

131. *Schechtman M.* 2004. Personality and Persistence: The Many Faces of Personal Survival // American Philosophical Quarterly. Vol. 41. № 2. P. 87–105.

132. *Schechtman M.* 2007. The Constitution of Selves. Ithaca: Cornell University Press.

133. *Schechtman M.* 2011. The Narrative Self // S. Gallagher (ed.) The Oxford Handbook of the Self. Oxford: Oxford University Press. P. 394–416.

134. *Schechtman M.* 2014. Staying Alive: Personal Identity, Practical Concerns, and the Unity of a Life. Oxford: Oxford University Press.

135. *Searle J.* 1994. The Rediscovery of Mind. Cambridge: MIT Press.

136. *Searle J.* 1997. The Mystery of Consciousness. N.Y.: New York Review of Books.

137. *Searle J.* 1998. Mind, Language and Society: Philosophy in the Real World. N.Y.: Basic Books.

138. *Searle J.* 1999. Intentionality: An Essay in the Philosophy of Mind. Cambridge: Cambridge University Press.

139. *Searle J.* 2003. Rationality in Action. Cambridge: Bradford Book.

140. *Searle J.* 2004. Mind: A Brief Introduction. Oxford: Oxford University Press.

141. *Searle J.* 2006. *Freedom and Neurobiology: Reflections on Free Will, Language, and Political Power.* N.Y.: Columbia University Press.

142. *Searle J.* 2007. *Biological Naturalism* // Max Velmans and Susan Schneider (eds.). *The Blackwell Companion to Consciousness.* Oxford: Blackwell Publishing Ltd. P. 325–334.

143. *Searle J.* 2008. *Philosophy in a New Century: Selected Essays.* Cambridge: Cambridge University Press.

144. *Sellars W.* 1991. *Science, Perception and Reality.* Atascadero: Ridgeview Pub. Co.

145. *Shoemaker D.* 2007. *Moral Address, Moral Responsibility, and the Boundaries of The Moral Community* // *Ethics.* Vol. 118. № 1. P. 70 – 108.

146. *Shoemaker D.* 2009. *Responsibility and Disability* // *Metaphilosophy.* Vol. 40. № 3–4. P. 438–461.

147. *Shoemaker D.* 2011a. *Moral Responsibility and the Self* // Gallagher S. (ed.). *The Oxford Handbook of the Self.* Oxford: Oxford University Press. P. 487–518.

148. *Shoemaker D.* 2011b. *Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility* // *Ethics.* Vol. 121. № 3. P. 602–632.

149. *Shoemaker D.* 2013. *Oxford Studies in Agency and Responsibility.* Volume 1. New York: Oxford University Press.

150. *Shoemaker D.* 2016. *Personal Identity and Ethics* // *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/>

archives/win2016/entries/identity-ethics/>.

151. *Shoemaker D., Tognazzini N.* 2014. *Oxford Studies in Agency and Responsibility*. Volume 2. New York: Oxford University Press.

152. *Shoemaker S.* 1970. *Persons and Their Pasts* // *American Philosophical Quarterly*. Vol. 7. № 4. P. 269–285.

153. *Shoemaker S.* 2001. *Realization and Mental Causation* // *Gillett C. and Loewer B. (eds.), Physicalism and its Discontents*. Cambridge, New York: Cambridge University Press. P. 74–98.

154. *Shoemaker S.* 2003. *Identity, Cause, and Mind*, Expanded Edition. Oxford: Clarendon Press.

155. *Smith A.* 2012. *Attributability, Answerability, and Accountability: in Defense of a Unified Account* // *Ethics*. Vol. 122. № 3. P. 575–589.

156. *Sosa E.* 1984. *Mind-Body Interaction and Supervenient Causation* // *Midwest Studies in Philosophy*. № 9. P. 271–281.

157. *Sosa E.* 1993. *Davidson's Tinking Causes* // *Heil J. and Mele A. (eds.), Mental causation*. Oxford: Clarendon Press. P. 41–50.

158. *Stoecker R.* 2010. *Davidson* // *O'Connor T. and Sandis C. (eds.), A Companion to the Philosophy of Action*. Oxford: Blackwell. P. 598 – 605.

159. *Strawson G.* 1994. *The Impossibility of Moral Responsibility* // *Philosophical Studies*. Vol. 75. № 1–2. P. 5–24.

160. *Strawson G.* 2005. *Against Narrativity* // *Strawson G. (ed.), The Self?* Oxford: Blackwell Publishing. P. 63–86.

161. *Strawson P.F.* 2008. Freedom and Resentment and Other Essays. London: Routledge.
162. *Swinburne R.* 1991. Personal Identity: the Dualist Theory // S. Shoemaker and R. Swinburne. Personal Identity. Oxford: Wiley-Blackwell. P. 1–65.
163. *Swinburne R.* 2013. Mind, Brain and Free Will. Oxford: Oxford University Press.
164. *Talbert M.* 2008. Blame and Responsiveness to Moral Reasons: Are Psychopaths Blameworthy? // Pacific Philosophical Quarterly. № 89. P. 516–535.
165. *Talbert M.* 2012. Accountability, Aliens, and Psychopaths: A Reply to Shoemaker // Ethics. Vol. 122. № 3. P. 562–574.
166. *Talbert M.* 2014. The Significance of Psychopathic Wrongdoing // Schramme T. (ed.). Being Amoral: Psychopathy and Moral Incapacity. Cambridge: MIT Press, P. 275–300.
167. *Talbert M.* 2016. Moral Responsibility: An Introduction (Key Concepts in Philosophy). Cambridge: Polity.
168. *Taylor C.* 1964. Explanation of Behavior. London: Routledge and Kegan Paul.
169. *Taylor M.* 1999. Imaginary Companions and the Children Who Create Them. New York, Oxford: Oxford University Press.
170. *Taylor M., Hodges S.D., Kohanyi A.* 2002–2003. The Illusion of Independent Agency: Do Adult Fiction Writers Experience Their Characters as Having Minds of Their Own? //

Imagination, Cognition, and Personality. Vol. 22. № 4. P. 361–380.

171. *Timpe K.* 2013. Free Will. Sourcehood and Its Alternatives. London, New York: Continuum International Publishing Group.

172. *van Inwagen P.* 1983. An Essay on Free Will. Oxford: Clarendon Press.

173. *van Inwagen P.* 2000. Free Will Remains a Mystery: The Eighth Philosophical Perspectives Lecture // *Noûs*, Vol. 34, Supplement: Philosophical Perspectives. Vol. 14. P. 1–19.

174. *van Inwagen P.* 2013. A Modal Argument for Incompatibilism // P. Russell and O. Deery (eds.). *The Philosophy of Free Will: Essential Readings From the Contemporary Debates*. Oxford, New York: Oxford University Press. P. 117–125.

175. *Vargas M.* 2012. Why the Luck Problem Isn't // *Philosophical Issues*. Vol. 22. № 1. P. 419–436.

176. *Vasilyev V.* 2009. «The Hard Problem of Consciousness» and Two Arguments for Interactionism // *Faith and Philosophy*. Vol. 26. № 5. P. 514–526.

177. *Velleman J.D.* 1996. Self to Self // *The Philosophical Review*. Vol. 105. № 1. P. 39–76.

178. *Vohs K., Schooler J.* 2008. The Value of Believing in Free Will: Encouraging a Belief in Determinism Increases Cheating // *Psychological Science*. Vol. 19. № 1. P. 49–54.

179. *Wallace R.* 1998. Responsibility and the Moral

Sentiments. Cambridge, MA: Harvard University Press.

180. *Wallace R.* 2011. Dispassionate Opprobrium: On Blame and the Reactive Sentiments // Wallace R.J., Kumar R., and Freeman S. (eds.). *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*. N.Y.: Oxford University Press, P. 348–372.

181. *Watson G.* 1987. Free Action and Free Will // *Mind*. Vol. 96. № 382. P. 145–172.

182. *Watson G.* 1996. Two Faces of Responsibility // *Philosophical Topics*. Vol. 24. № 2. P. 227–248.

183. *Watson G.* 2004. Responsibility and the Limits of Evil: Variations on a Strawsonian Theme // *Watson G. Agency and Answerability: Selected Essays*. Oxford: Oxford University Press. P. 219–259.

184. *Watson G.* 2011. The Trouble with Psychopaths // Wallace R.J., Kumar R., and Freeman S. (eds.). *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*. N.Y.: Oxford University Press. P. 307–331.

185. *Wegner D.* 2002. *The Illusion of Conscious Will*. Cambridge, London: Bradford Books, The MIT Press.

186. *Wegner D.* 2004. Précis of *The Illusion of Conscious Will* // *Behavioral and Brain Sciences*. № 27. P. 649–692.

187. *Williams B.* 1956–1957. Personal Identity and Individuation // *Proceedings of the Aristotelian Society: New Series*. Vol. 57. P. 229–252.

188. *Williams B.* 2008. *Shame and Necessity*. Los Angeles:

University of California Press.

189. *Wolf S.* 1981. The Importance of Free Will // *Mind*. № 90. P. 366–378.

190. *Wolf S.* 1993. *Freedom within Reason*. N.Y.: Oxford University Press.

191. *Yablo S.* 1992. Mental Causation // *The Philosophical Review*. Vol. 101. № 2. P. 245–280.

192. *Zimmerman M.* 1988. *An Essay on Moral Responsibility*. Totowa, NJ: Rowman and Littlefield.

193. *Zimmerman M.* 2010. Responsibility, Reaction, and Value // *Journal of Ethics*. Vol. 14. № 2. P. 103–115.