

НА ПИКАХ АНТЕНН

Д.Б. Волков

ЧТО ДОКАЗЫВАЮТ МАНИПУЛЯЦИИ С «АРГУМЕНТОМ МАНИПУЛЯЦИИ» Д. ПЕРЕБУМА?

Аннотация. Предметом исследования в данной статье является проблема свободы воли. Автор статьи представляет критику одного из наиболее важных современных аргументов против компатибилизма в вопросе о свободе воли – «Аргумента Манипуляций» Д. Перебума. Аргумент Манипуляций построен на аналогии между внешней манипуляцией действиями агента и каузальной детерминацией. На основе этой аналогии Перебум делает вывод об отсутствии моральной ответственности агентов, действующих в условиях каузальной детерминации. Автор данной статьи предлагает расширение мысленного эксперимента Перебума. В ходе этого исследования автор приходит к выводу о том, что Аргумент Манипуляций не доказывает несовместимость детерминизма с моральной ответственностью. По мнению автора, аргумент Перебума показывает только необходимость существования целостной личности и последовательной личностной истории для моральной ответственности. Методом данного исследования является концептуальный анализ, характерный в целом для аналитической философии. Автор также использует мысленные эксперименты для иллюстрации и аргументации своей позиции. Основным выводом данного исследования является утверждение ложности аргумента манипуляций Перебума. Автор статьи показывает, что аргумент манипуляций не может служить основанием для утверждения несовместимости детерминизма с моральной ответственностью. По мнению автора, аргумент Перебума показывает только необходимость существования целостной личности и последовательной личностной истории для моральной ответственности.

Ключевые слова: свобода воли, моральная ответственность, Д. Перебум, компатибилизм, Аргумент Манипуляций, Аргумент четырёх случаев, детерминизм, философия свободы, инкомпатибилизм, М. Маккена.

Review. The subject under research of the present article is the problem of free will. The author of the article presents a critical review of one of the most important modern arguments against compatibilism in the free will debate, Derk Pereboom's 'Manipulation Argument'. The Manipulation Argument is built upon the analogy between external manipulation of the agent's actions and causal determination. Based on that analogy, Pereboom concludes that agents who act in a situation of causal determination do not feel moral responsibility. The author of the present article suggests that we should extend the borders of Pereboom's mental experiment. In the course of his research the author comes to the conclusion that the Manipulation Argument does not prove the incompatibility of determinism and moral responsibility. In the author's opinion, Pereboom's argument only proves that moral responsibility requires an integrated personality and sequential personal history. The method of the present research is the conceptual analysis mostly used in analytical philosophy. The author also conducts mental experiments to illustrate and prove his point of view. The main conclusion of the present research is the author's proof of the failure of Peterboom's manipulation argument. The author of the article shows that the manipulation argument does not prove the statement that determinism and moral responsibility are incompatible. According to the author, Pereboom's argument only proves that moral responsibility requires an integrated personality and sequential personal history.

Keywords: free will, freedom of the will, moral responsibility, Derk Pereboom, compatibilism, Manipulation Argument, four-case argument, determinism, philosophy of freedom, incompatibilism, Michael McKenna.

Часто за свои действия люди несут ответственность, моральную и иногда юридическую. На основании этой ответственности мы благодарим друг друга, награждаем, наказываем, обижаемся. Но как возможна мораль-

ная ответственность в мире, где все действия каузально обусловлены, где каждое событие имеет причину, и цепочка причин уходит в бесконечное прошлое? Уместны ли такие реакции в случае, если действия агентов предопределены? Эти вопросы

традиционно являются предметом исследования философии. Однако в последние десятилетия, в особенности в связи с развитием нейронаук, они становятся ещё более актуальными. В ходе эмпирических исследований проясняются механизмы, управляющие поведением людей. Это даёт благодатную почву для философских изысканий. Новые эмпирические возможности позволяют философам совершенствовать мысленные эксперименты и обосновывать свою позицию. Одним из таких усовершенствованных мысленных экспериментов является «Аргумент манипуляций» Д. Перебума. Этот аргумент представляет собой новую атаку на компатибилизм – тезис о том, что моральная ответственность и свобода воли совместима с детерминизмом. Американский философ занимает инкомпатибилистскую позицию. Он считает, что свобода воли не совместима как с индетерминизмом, так и с детерминизмом. Перебум утверждает, что свобода воли отсутствует. Это любопытная позиция. Её придерживаются не более 15 % англо-язычных философов [1, с. 465-600]. Но она, безусловно, заслуживает внимания. В этой статье я попытаюсь разобраться с основным аргументом в пользу инкомпатибилизма Перебума – «Аргументом Манипуляции» и попытаюсь определить, доказывает ли этот аргумент несовместимость моральной ответственности, свободы воли и детерминизма.

«Аргумент манипуляций» Перебума (этот аргумент ещё называют Аргументом четырёх случаев) построен на интуиции о том, что действия в результате манипуляций не являются ответственными действиями. По мнению Перебума, манипулирование извне и каузальная 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, свободы и, следовательно, моральной ответственности. Перебум начинает исследовать интуиции с умозрительного, но очевидного, по его мнению, случая, где агент является марионеткой в руках других агентов – ученых. В этой вымышленной ситуации ученые управляют каждым действием агента, и потому агент, кажется, не несет ответственности. Далее Перебум приводит другие случаи манипуляций, все более и более приближенные 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, обыкновенных условиях. «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 ответственности. Во всех ситуациях профессор Плам совершает убийство Уайта из корыстных соображений. Итак, первая ситуация – случай, где Плам является марионеткой ученых:

Сценарий 1. Радиоуправляемый агент. Команда нейроученых с помощью радиопередатчиков управляют мозгом и, соответственно, ментальными состояниями профессора 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 в целом склонностям его характера, так как его характер в целом располагает к эгоистическим соображениям. Желание совершить убийство не противоречит более высокопорядковым желаниям Плама. Т.е. 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 [6, с. 424].

По замыслу Перебума радиоуправляемый агент должен представляться читателям как агент лишённый моральной ответственности. Ведь его действия – результат манипуляций. При этом манипулируемый Плам по основным параметрам удовлетворяет критериям моральной ответственности, выдвигаемым компатибилистами. В частности он удовлетворяет критериям, которые защищают Юм, Айер, Франкфурт, Уоллас, Фишер и Равицца. Это должно продемонстрировать, что эти компатибилистские теории не предлагают достаточных условий для 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 агент, кроме одного обстоятельства. В начале его жизни команда нейрочученых запрограммировала его мозг так, что его рассуждения принимаю крайне эгоистический характер. Неизбежным следствием этих рассуждений является решение Плама убить Уайта. Нейрофизиологическая реализация и история состояний Плама полностью соответствует Сценарию 1. Эгоистическое решение Плама не противоречит в целом склонностям его характера. Желание совершить убийство не противоречит 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 (и он практически дублирует приведенный выше пример с Марио-неткой). Все выдвинутые ранее компатибилистские критерии свободы воли удовлетворены. А интуиции, тем не менее, склоняют к отрицанию 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 во времени, когда происходит манипуляция. Радиоуправляемым агентом ученые манипулирует в реальном времени. А запрограммированный агент изначально создан так, чтобы совершить преступление. Но эти различия в отношении моральной ответственности, по мнению Перебума, ничего не меняют. Как, впрочем, ничего не должны менять и различия между вторым и третьим Сценарием.

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 склонно-

сти не противоречат в целом характеру Плама. И у него нет непреодолимых желаний.

Третий сценарий ещё ближе к реальным условиям. Он вполне соответствует ситуации, которая по представлениям некоторых утопистов, может сложиться в обществе будущего. Предположительно, он близок и реальной ситуации, в которой воспитывались мальчики из аристократических семейств в XIV и XV вв. 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. Т.е. 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 происходит в мире каузально детерминировано событиями прошлого и законами природы. Плам – это обычный человек, появившийся и воспитанный в обычных условиях. Его рассуждения часто (но не всегда) носят крайне эгоистический характер. Его решение убить Уайта происходит из склонностей его характера, в ходе рациональных рассуждений. Это его желание полностью соответствует его второпорядковому желанию, иметь первопорядковое желание. Нейрофизиологическая реализация и история состояний Плама полностью соответствует Сценарию 1-3. Плам способен в целом регулировать своё поведение моральными принципами и не действует из непреодолимых желаний.

Если рассуждать последовательно, то и в четвертом случае моральная ответственность отсутствует. Ведь этот случай во многом аналогичен предыдущим. Сценарии подобраны таким образом, что каждый следующий только незначительно отличается от предыдущего. Перебум сначала представляет оригинальный контрпример на компатибилистские критерии 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 четырех случаях, снимает ответственность с агента, т.к. она находится вне зоны его контроля.

Агент не может контролировать не предшествующие события, ни законы природы, и именно поэтому не несет ответственности за производимые действия. Их этого следует, что не только ученные компатибилистские теории не могут предоставить достаточных условий для свободы воли и ответственности. Но и любые компатибилистские теории в принципе не могут представить таких критериев. Свобода воли и моральная ответственность не совместима с каузальной детерминацией. Компатибилизм ложен.

Аргумент Перебума можно представить в более формальном виде. Для удобства введем аббревиатуру ККС – компатибилистические критерии структуры свободного агента. Эта аббревиатура будет обозначать все необходимые условия, которые компатибилистские теории представляют для классификации действия как свободного действия агента. Эти критерии входят соответствие действия склонностям характера, отсутствие непреодолимых желаний, соответствие первопорядковых желаний второпорядковым желаниям, согласованность желаний с рациональными мотивами агента. С использованием 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 с удовлетворением ККС, не действует по своей воле, и потому не несет моральной ответственности ($D \rightarrow \neg R$)

Первая посылка интуитивно достоверна. 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 достоверным. Аргумент кажется изобретательным и достоверным. Означает ли это, что Перебуму удалось опровергнуть компатибилистические теории?

Аргумент Манипуляций – панель с шестью тумблерами

Чтобы проверить Аргумент Манипуляций нужно разъяснить некоторые моменты. Во-первых, нужно убедиться в правомерности обобщения всех случа-

ев 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, нужно поочередно вносить в них изменения и «смотреть», что получится. Определим, для начала, какими параметрами можно варьировать.

В Сценариях аргумента Манипуляций есть несколько ключевых параметров. На первый взгляд это: (1) тип манипуляторов; (2) время манипуляции; (3) компатибилистические критерии свободного агента; (4) характер действия; (5) последовательность сценариев; (6) 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 включая 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, 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 обречён на неудачу с самого начала, ведь даже, если ответственность действи-

тельно отсутствует в первых трёх случаях, её обобщение на четвёртый не состоятельно. Вот убираем агентов из первых трех сценариев, выключаем тумблер «манипуляторы» и... видим, 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 остаётся. Плам совершит убийство, но моральной ответственности за него нести не должен. В целом можно согласиться с Перебумом. Интуиции о наличии моральной ответственности, скорее всего, не возникнут. Хотя в этом случае ситуация уже не так однозначна. В любом случае, очевидно, что причина не в том, что Пламом управляют другие агенты, точнее, не только в том, но и, очевидно, в чем-то другом.

Другой параметр – *время манипуляции* (2). В первом случае Плам управляется в режиме реального времени, т.е. момент манипуляции непосредственно предшествует действию. В другом случае манипуляция совершается в детстве Плама, за много лет до совершения убийства (Сценарий 3). А ещё в одном Сценарии воздействие происходит намного раньше, до существования Плама. Четвёртый сценарий предполагает «манипуляцию» в бесконечно далеком прошлом: в момент, когда законы природы были сформированы. Т.е. манипуляция может отодвигаться на сколь угодно длительное время от действия. Но влияет ли 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, играть значения. Плам всё равно не несет ответственность за действия предопределенные в далеком прошлом. Выходит, тумблер «время» не играет принципиального значения. Хотя в этом случае у меня интуиция уже не так однозначна. В любом случае, очевидно, что причина не в том, что Пламом управляют в режиме реального времени, точнее, не только в этом, но и в чем-то другом.

Возможно, тогда дело в *конкретных компатибилистских критериях* (3) свободного ответственного действия? Попытаемся управлять этим параметром. 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 полностью быть удовлетворены, и при этом моральной ответственности нет. Может, тогда стоит пополнить или изменить их список? По мнению Перебума, и даже некоторых компатибилистов, пополнение ККС не изменит принципиально интуиций. Почему? Допустим, существует «лучшая компатибилистская теория», теория в которой учтено условие X, являющееся лучшим кандидатом на достаточность для свободного ответственного действия. Все равно, кажется возможным модифицировать все Сценарии Аргумента Манипуляций так, чтобы они учитывали это условие X, при этом ответственность не будет возникать. Просто 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 философские произведения выглядят как гибриды интеллектуальной прозы и жестоких второсортных детективов. Объяснение у этого простое. Философам нужны примеры, которые вызывали бы однозначную реакцию. В этом смысле убийства подходят. Но что произойдет, если убийство Уайта заменить на другое действие, действие, заслуживающее одобрения и поощрения? Предлагаю новый сценарий с другим типом действия.

Сценарий 5. Радиоуправляемый герой. Команда нейроученых с помощью радиопередатчиков управляют ментальными состояниями Плама. Плам видит, что из окон соседнего дома идет дым. От соседей он узнает, что в горящем доме 4-х-летний ребёнок. Перед тем, как Плам начинает размышлять, нейроученые 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. Спасённый мальчик обнимает Плама, ро-

дители благодарят его и восхищены мужественным поступком. Плам отказывается принимать благодарности и скромно прощается с новыми знакомыми. Он чуть сам не погиб, спасая ребенка, но если бы аналогичный случай произошел ещё раз, Плам бы поступил точно так же.

Во всех остальных вещах этот сценарий похож на Сценарий 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. Несёт ли альтруистичный Плам моральную ответственность? Заслуживает 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 разные в отношении Сценариев с положительными и отрицательными поступками. Да, характер Плама стал более альтруистическим 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 С. Вульф и Д. Нелкин. Это странная ситуация. Выходит, изменение характера поступка даже в случаях с манипуляциями влияет на формирование моральной ответственности? Но к какому выводу это нас толкает? К убеждению, что иммунитетом в отношении аргумента манипуляций обладают только положительные поступки и 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 ответ. «Я согласен, что интуитивно более уместно хвалить Плама в этом примере [с положительным поступком – Д.В.], чем в исходном Сценарии... Но для меня не очевидно, что эта похвала и моральное одобрение являются признаками моральной ответственности в значении простого воздаяния [basic desert – Д.В.]...» [7, с. 102]. Людям в принципе свойственно хвалить других за врожденный интеллект, внешнюю привлекательность, таланты. Хотя эти качества во-

все не заслуга агентов. Хвала в данном случае, по мнению философа, выражение восторга, радости, позитивного отношения, а не признак моральной ответственности. Возможно, и в случае с модифицированным сценарием реакция на поступок Плама просто выражение радости в отношении спасенного ребенка. Другими словами Перебум считает, что этот тумблер не является искомым, он не влияет на лампочку моральной ответственности. Это ложноположительный результат. Допустим. Тогда переключим следующий тумблер.

Поэкспериментируем с ещё 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 решат, что в этом случае, моральная ответственность существует. Такое отношение к случаю исходит из нормальной практики человеческих отношений. Но даже более скептическая позиция приемлема. Допустимо считать, что в ситуации с детерминированным агентом не ясно, является ли этот агент свободным, или нет, и, соответственно, несет ли он ответственность. Теперь развернём 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 МакКенна. Его вывод нейтрален: «Я пытаюсь продемонстрировать, что из Сценариев... не следует, что Плам не свободен и не имеет моральной ответственности» [5, с. 153]. Только из одного детерминизма вывести отсутствие свободы воли и ответственности нельзя. Но если тактика обобщения верна, следовательно, агностический вывод достоверен. Мы не можем знать, какой тумблер влияет на состояние лампочки. Но, более того, мы даже не можем знать вообще, что происходит с лампочкой как в случае выключения всех тумблеров, так и по-отдельности.

Однако в чём 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 твёрдого 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 убеждённого агностика. Сторонники этой позиции считают, что в отношении моральной ответственности не ясно, совместима она с детерминизмом или нет. Но убеждённые агностики не видят возможности эту позицию дальше прояснить. Для них вопрос закрыт на этом. Можно также начинать с позиций убеждённого компатибилиста или убеждённого инкомпатибилиста. Но и эти позиции не предполагают продуктивной дискуссии. Более продуктивны открытые исследовательские позиции, ищущие согласование с новыми доводами и фактами. Перебум считает, что агностический вывод из аргумента МакКенны следует только из позиции убеждённого компатибилиста, закрытого для принятия новых фактов. А новые факты здесь есть значение. Это факты 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, т.к. они осведомлены о своих желаниях и действиях, но не знают о причинах, их вызывающих. Именно эти новые факты, факты о детерминирующих обстоятельствах, должны склонить чаши весов в сторону инкомпатибилизма. И это как раз 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 перекинуть последний, шестой тумблер.

Это *тип детерминации (б)*. Напомню, во всех Сценариях аргумента причины с необходимостью определяют результат, т.е. для агента 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 только с вероятностью определяют его действия, интуиции в отношении его ответственности не будут поколеблены. В одном случае манипуляторы управляют со стопроцентным успехом, в другом – только с некоторой вероятностью – для Плама это ничего не меняет.

Это любопытная ситуация. Она, казалось бы, подрывает аргументацию Перебума (соображения подобного рода против Перебума использует МакКенна). 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, стандартом 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 ещё 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 в случаях с каузальной детерминацией, а также и в случаях с индетерминированными событиями. Для того, чтобы понять соотношение между двумя этими причинами он приводит такую аналогию. Предположим, где-то прорвало плотину. Инженера спрашивают, в чем причина. Инженер объясняет, что плотину прорвало, так как резервуар была наполнен более, чем одним миллионом тонн воды. На это его начальник возражает: вода здесь не при чем. Представь,

что резервуар был бы наполнен более, чем одним 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, даже если резервуар переполнился бы нефтью. Ведь есть ещё более глубокая, общая причина прорыва: превышение давления жидкости на стену плотины. Прорыв плотины, как и отсутствие моральной ответственности, имеет причины разных уровней общности и глубины. Таким образом, в описанных ситуациях причиной является наличие каузальной детерминации, в других ситуациях – мог быть индетерминизм, но в любом случае общей причиной является отсутствие необходимого контроля.

Возражение на Аргумент Манипуляции. Седьмой тумблер

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 (можно также сказать, что проблема в приписывании поступка. Когда этот поступок можно приписать Пламу в силу его личностных характеристик, он несет ответственность. Когда этот поступок ему нельзя приписать, а, скорее, его можно приписать ученым, управляющим экспериментом, Плам ответственности не несет). **Интуиция о моральной ответственности возникает тогда, когда мы представляем Плама единой целостной личностью, проживающую последовательную, достаточно полную и целенаправленную историю, и исчезает, когда этой целостности и последовательности нет.** На мысль об этом наводит уже проделанный эксперимент с типом поступка. Напомню, Перебум говорит, что благодарность в случае альтруистического поступка, это не свидетельство интуиции 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, его поступок, а не обстоятель-

ства ситуации. И это показатель наличия моральной ответственности в смысле «воздаяния». Здесь я согласен с Вульф и Нелкин.

Но мне кажется неверным их интерпретация причины. Эти философы предполагают, что моральная ответственность асимметрична. Т.е. 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 – существует. По моему мнению, оснований для такого экстравагантного решения, утверждения асимметрии, не достаточно. 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, чтобы совершить альтруистический поступок. Решение согласовано с его желаниями, его характерам (пусть даже «новым»), жизненной историей и последствиями. 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 Плама. 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?

Случай *Альтруистический радиоуправляемый агент* отличается от других случаев тем, что читатель представляет, как агент переживает последствия своих действий. Плам рискует, спасая ребенка, и сам чуть не погибает. Это позволяет выявить, какого рода личностью он является. Во всех остальных Сценариях, эти последствия скрыты. В Сценариях, описываемых Перебумом, нет явного указания на переживание последствий агентом. Нет возможности наблюдать характер агента, нет ощущения целенаправленности его взаимодействия с внешним миром и его соответствия стандарту рациональности. 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 ещё одно преступление. Ему не нужны свидетели и конкуренты. Это происходит, конечно, без каких-то дополнительных манипуляций со стороны нейроученых. Они ведь ещё до первого убийства сдвинули его шкалу ценностей в сторону крайней эгоистичности. Дальше всё 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 ей оставаться. Его поступки сообразны его личным качествам, его характеру и ценностям. Их непосредственно относятся к Пламу. Разве по справедливости он не должен отвечать за преступления? Другое дело, если бы после убийства Уайта, Плам оказался бы морально раздавленным. Если бы он винил себя, отчаянно пытался разобраться, как все это произошло, и отправился 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 лишили Плама моральной ответственности 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 продолжают, 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 лагун, допущенных в мысленном эксперименте. Именно эти лагуны не позволяют судить о Пламе как о самостоятельной ответственной личности. А ведь для моральной ответственности нужен не только виновный поступок 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.

Список литературы:

1. Bourget D., Chalmers D. What Do Philosophers Believe? // Philosophical studies. 2014. No. 170. P. 465-600.
2. Fischer J., Kane R., Pereboom D., Vargas M. Four views on free will. Malden, 2007.
3. Kane R. The significance of free will. Oxford, 1998.
4. Kane R. Free will: New directions for an ancient problem // Free will / Ed. by R. Kane. Malden, 2002.
5. McKenna M. A Hard-line reply to Pereboom's four-case manipulation Argument // Philosophy and Phenomenological Research. July 2008. Vol. LXXVII. No. 1. P. 153.
6. Pereboom D. Optimistic Skepticism about Free Will. The philosophy of Free Will / Ed. by P. Russel and O. Deery. P. 424.
7. Pereboom D. Free Will, Agency, and Meaning in Life. Oxford, 2014. P. 102.
8. Mele A. A critique of Pereboom's 'four-case argument' for incompatibilism // Analysis. 2005. No. 65(285). P. 75-80.
9. Pereboom D. Living without Free Will, Cambridge: Cambridge University Press, 2001.
10. Pereboom D. Reasons Responsiveness, Alternative Possibilities, and Manipulation Arguments Against Compatibilism; Reflections on John Martin Fischer's My Way // Philosophical Books. 2006. No. 47. P. 198-212.
11. Pereboom D. A Hard-Line Reply to the Multiple-Case Manipulation Argument // Philosophy and Phenomenological Research. 2008. No. 77. P. 160-170.
12. Pereboom D. Responses to John Martin Fischer and Dana Nelkin (on my Free Will, Agency, and Meaning in Life) // Science, Religion and Culture. 2014. No. 1. P. 218-225.

13. Pereboom D. Free Will in The Oxford Handbook of the History of Ethics, Roger Crisp, ed., Oxford: Oxford University Press, 2013. P. 606-737.

References (transliteration):

1. Bourget D., Chalmers D. What Do Philosophers Believe? // Philosophical studies. 2014. No. 170. P. 465-600.
2. Fischer J., Kane R., Pereboom D., Vargas M. Four views on free will. Malden, 2007.
3. Kane R. The significance of free will. Oxford, 1998.
4. Kane R. Free will: New directions for an ancient problem // Free will / Ed. by R. Kane. Malden, 2002.
5. McKenna M. A Hard-line reply to Pereboom's four-case manipulation Argument // Philosophy and Phenomenological Research. July 2008. Vol. LXXVII. No. 1. P. 153.
6. Pereboom D. Optimistic Skepticism about Free Will. The philosophy of Free Will / Ed. by P. Russel and O. Deery. P. 424.
7. Pereboom D. Free Will, Agency, and Meaning in Life. Oxford, 2014. P. 102.
8. Mele A. A critique of Pereboom's 'four-case argument' for incompatibilism // Analysis. 2005. No. 65(285). P. 75-80.
9. Pereboom D. Living without Free Will, Cambridge: Cambridge University Press, 2001.
10. Pereboom D. Reasons Responsiveness, Alternative Possibilities, and Manipulation Arguments Against Compatibilism; Reflections on John Martin Fischer's My Way // Philosophical Books. 2006. No. 47. P. 198-212.
11. Pereboom D. A Hard-Line Reply to the Multiple-Case Manipulation Argument // Philosophy and Phenomenological Research. 2008. No. 77. P. 160-170.
12. Pereboom D. Responses to John Martin Fischer and Dana Nelkin (on my Free Will, Agency, and Meaning in Life) // Science, Religion and Culture. 2014. No. 1. P. 218-225.
13. Pereboom D. Free Will in The Oxford Handbook of the History of Ethics, Roger Crisp, ed., Oxford: Oxford University Press, 2013. P. 606-737.