

# «Тезис о сознании» и моральная ответственность в исследованиях Нила Леви

ДМИТРИЙ ВОЛКОВ

Содиректор, НП «Московский центр исследования сознания».

Адрес: 119121, Москва, ул. Бурденко, 14А. E-mail: [dvolkoff@gmail.com](mailto:dvolkoff@gmail.com).

*Ключевые слова:* свобода воли; моральная ответственность; тезис о сознании; сознание; Нил Леви.

В статье проанализирован «тезис о сознании» Нила Леви, состоящий в том, что сознание необходимо для моральной ответственности, и аргументы в его пользу. Под термином «сознание» имеется в виду информационный процесс в мозге, который влияет на поведение агента и о котором он может представить вербальный отчет. Противниками «тезиса о сознании» являются философы-экспрессивисты Номи Арпэли, Анжела Смит, Джордж Шер. Аргумент Леви построен на теории сознания как «глобального рабочего пространства». Леви считает, что только сознание обеспечивает интеграцию информации на персональном уровне и обуславливает возможность гибкого поведения агента. Автоматизмы и поступки без осознания моральных аспектов не интегрируются на персональном уровне. В свою очередь интеграция информации и гибкий контроль являются основанием для ответственности согласно двум наиболее распространенным теориям: «теории самовыражения реального „Я“» и «теории контроля». Следовательно, «тезис о сознании» является доказанным.

Автор данной статьи считает, что наиболее уязвимой стороной аргумента Леви является его определение сознания. «Тезис о сознании», скорее всего, является спорным в контексте двух наиболее распространенных теорий сознания: теории «квалофилов» и «иллюзионистов». Квалофилы считают, что семантическое содержание определяется качественными состояниями, и интеграция информационных процессов сама по себе не дает основания для формирования убеждений на персональном уровне. Следовательно, сознание, как его определяет Леви, для квалофилов не должно представляться необходимым условием для ответственности. Иллюзионисты считают, что убеждения приписываются агентам извне при интерпретации их поведения. Поэтому даже те информационные процессы, о которых субъект не может легко представить отчет, выражают убеждения личности. Из этого также следует, что сознание, как его определяет Леви, не должно быть необходимым условием ответственности.

**П**РОБЛЕМА сознания — одна из ключевых для аналитической философии. Современные философы совместно с нейрочеными пытаются определить, каким образом состояния мозга соотносятся с ментальными, каковы корреляты сознания и какую роль оно играет в поведении человека. Самостоятельными направлениями исследований в аналитической традиции являются проблемы свободы воли и моральной ответственности. Философы спорят о том, насколько свобода воли, лежащая в основании моральной ответственности, совместима с каузальным детерминизмом — утверждением о том, что все происходящие события с необходимостью предопределены причинами, уходящими в далекое прошлое. Чаще всего дискуссии о сознании и моральной ответственности не пересекаются, что придает особую значимость исследованиям соотношения этих категорий Нила Леви. Леви пытается доказать, что сознание является необходимым условием моральной ответственности. Это ключевое утверждение автор называет «тезисом о сознании». Его доказательству посвящены новейшие работы Леви, в частности монография «Сознание и моральная ответственность»<sup>1</sup>.

Путь, которым Леви приходит к такому выводу, был несколько неожиданным. Извилистой оказалась и его академическая карьера. Философ защитил не одну, а две докторские диссертации. Первая, датируемая 1995 годом, была посвящена континентальной философии: экзистенциализму, постмодернизму, работам Сартра и Фуко. Вторая, посвященная проблеме моральной ответственности, написана уже в аналитической традиции, как и самая известная книга автора — «Нейроэтика: вызовы XXI веку»<sup>2</sup>. Эта монография любопытна тем, что фактически открыла новую междисциплинарную область — нейроэтику, которая направлена на прояснение связи между нейрофизиологическими представлениями о мозге и философскими теориями о свободе воли и основах морали. В отдельной главе «Нейроэтики» рассмотрено от-

1. См.: *Levy N. Consciousness and Moral Responsibility*. Oxford: Oxford University Press, 2014.

2. См.: *Idem. Neuroethics: Challenges for the 21st Century*. Cambridge, UK: Cambridge University Press, 2007.

ношение сознания, свободы воли и моральной ответственности. Уже здесь автор делает предварительный вывод о том, что сознание, скорее всего, необходимо для моральной ответственности.

Следующая книга Леви — «Жестокая случайность: как случайность подрывает свободу воли и моральную ответственность»<sup>3</sup> — получилась очень провокационной и оригинальной, поскольку отрицала возможность существования свободы воли и моральной ответственности, а также рациональных оснований для моральных оценок — одобрения и осуждения. Причину отсутствия ответственности Леви усматривал не в детерминизме, подобно большинству инкомпатибилистов, а в случайности. Леви выделил два вида случайности: конституирующую (*constitutive*) и ситуационную (*present*). По его мнению, перед исследователями стоит дилемма: либо допускать, что действия агентов являются результатом наследственности, либо нет. Если это результат наследственности, то агент является жертвой конституирующей случайности, а если нет — ситуационной. В любом случае случайность лишает агента контроля как решающего условия свободы и моральной ответственности. Следовательно, никто не несет ответственности за свои действия. Представляется, впрочем, что такой вывод плохо согласуется с идеями «Нейроэтики» и новой монографии Леви. Зачем изучать связь между моральной ответственностью и сознанием, если ответственности не существует?

В предисловии к «Сознанию и моральной ответственности» автор разъясняет, как согласуются выводы этой работы с предыдущими. Во-первых, изучение некоторых неочевидных аспектов общераспространенных представлений оправдано само по себе. Во-вторых, аргументы о полном отсутствии ответственности не всем кажутся убедительными. И если в этой книге удастся доказать, что для ответственности необходимо сознание, то, по крайней мере, будет считаться, что часть людей не несет ответственности за совершенные поступки. Это, по мнению Леви, уже частичная победа. Кроме того, тезис о сознании, с точки зрения автора, позволяет определить, какие действия могут быть морально оценены и какие наказания оправданны. Данные аспекты не связаны с проблемой моральной ответственности напрямую, ведь наказания и оценки могут выноситься независимо от ответственности — например, из соображений, связанных с регулированием общественной жизни. Представленные в новой моногра-

3. См.: *Idem*. *Hard Luck: How Luck Undermines Free Will and Moral Responsibility*. Oxford: Oxford University Press, 2015.

фии идеи, таким образом, не противоречат предыдущим выводам Леви, но отчасти лежат в русле его исследований и подкрепляют более ранние полемические утверждения. Но с кем он ведет эту полемику? Кто не готов принять «тезис о сознании»?

Еще в «Нейроэтике» Леви замечает, что выводы некоторых ученых о психологии принятия решений могут восприниматься как критика его «тезиса о сознании». Речь идет об исследованиях Бенджамина Либета<sup>4</sup> и Дэниела Вегнера<sup>5</sup>. Эксперимент Либета, направленный на то, чтобы раскрыть механизм принятия решений, пожалуй, самый известный в современной нейрофизиологии. Во время эксперимента испытуемому предлагалось двигать рукой при возникновении произвольного желания. При этом сам испытуемый должен был фиксировать время появления такого желания. Параллельно с такой субъективной оценкой испытуемого Либет с помощью датчиков измерял мозговую активность, а именно «потенциал готовности», который всегда служит сигналом к действию. Опыт позволил Либету установить, что фиксируемое испытуемым время появления желания отстает от момента регистрации/возникновения «потенциала готовности». Иными словами, бессознательные процессы в мозге, определяющие действие, предшествуют сознательному желанию. Таким образом, сознательные состояния лишь регистрируют свершившийся факт, запаздывая относительно нейрофизиологического сигнала к действию.

К выводу о том, что сознание лишь регистрирует решения, которые принимаются бессознательно, пришел и Вегнер. Он обратил внимание на феномен диссоциации воспринимаемого и фактического контроля. В качестве примеров диссоциации он привел лабораторные эксперименты, в частности опыт с «чужими руками». В этом эксперименте на испытуемого надевают робу и ставят его перед зеркалом. Помощник встает сзади и просовывает руки в перчатках в рукава робы так, чтобы испытуемому казалось, будто в зеркале отражаются его руки. Ведущий инструктирует, что нужно сделать: хлопнуть в ладоши, показать кулак и т. д., а помощник выполняет эти действия. В ходе эксперимента испытуемому начинает казаться, что команды выполняет он сам, своими собственными руками, в чем и проявляется феномен диссоциации. На основании подобных наблюдений Вегнер делает вывод

4. См.: *Libet B. Mind Time: Temporal Factors of Consciousness*. Cambridge, MA: Harvard University Press, 2005.

5. См.: *Wegner D. The Illusions of Conscious Will*. Cambridge, MA: MIT Press, 2003.

о том, что ощущение волевого контроля и наличие фактического не совпадают, то есть сознательная воля — это иллюзия.

Может показаться, что открытия Либета и Вегнера опровергают тезис Леви, но это не так. Даже если признать их эксперименты и интерпретации правдоподобными, они не ставят под сомнение «тезис о сознании», как его определяет автор. Вместо этого они подрывают другие тезисы: например, о том, что действие инициируется полностью сознательно, или что сознание играет автономную каузальную роль, или что сознательная личность является независимым источником действия. По мнению Леви, эти тезисы ошибочны, независимо от результатов конкретных эмпирических исследований. Сознательные состояния вплетены в каузальные цепочки нейронных состояний, большая часть которых бессознательна. Достаточно «кабинетной философии», чтобы сделать вывод о том, что сознание не может быть автономным. Сознание фактически не играет той роли в действии, которую Либет и Вегнер считают необходимой для существования моральной ответственности<sup>6</sup>. Сознательные процессы так или иначе каузально зависимы от бессознательных. Но из этого не следует отсутствие, во-первых, возможности моральной ответственности и, во-вторых, необходимости сознания для моральной ответственности. По мнению Леви, сознанию просто отведена иная роль — интеграция информации на личностном уровне.

Традиционно сознательные состояния определяют как имеющие феноменальный, качественный аспект. Эта традиция опирается на важное различие, проведенное Дэвидом Чалмерсом, который поместил его в основание так называемой трудной проблемы (*hard problem*). «Трудная проблема» состоит в объяснении того, как физические состояния мозга соотносятся с феноменальными субъективными состояниями. Однако эта проблема лежит в стороне от исследований Леви. Когда автор говорит о сознании, он имеет в виду не феноменальные состояния, а информационные.

Те сознательные состояния, о которых я говорю, — это состояния с информационным содержанием; имеют ли они феноменальное содержание или нет... не имеет значения для наших целей<sup>7</sup>.

Под информационными философ понимает такие состояния, в которых агент отдает себе отчет и по поводу которых обычно может высказываться. Таким образом, сознание, в определении

6. *Levy N. Neuroethics*. P. 231.

7. *Idem. Consciousness and Moral Responsibility*. P. 29.

Леви, синонимично осведомленности или информированности (*awareness*). Критериями такой осведомленности служат влияние этих ментальных содержаний на поведение и возможность формировать вербальные отчеты о наличии этой информации при неспецифических внешних подсказках. В качестве иллюстраций Леви приводит несколько примеров.

Предположим, Эмили забыла о годовщине собственной свадьбы, поэтому вместо романтического ужина она собирается на вечеринку с подругами. Тем не менее информация о том, что сегодня годовщина свадьбы, влияет на ее поведение: она надевает кольцо, подаренное мужем на предыдущую годовщину именно в силу тех обстоятельств, при которых оно было подарено. Можно ли в данном случае назвать информацию о годовщине осознанной? По мнению Леви, предложенное описание не подходит под названную категорию. Информация влияет на поведение Эмили, но, если она не может ответить, почему надела кольцо, информация не может быть извлечена достаточно легко, то есть, по определению Леви, осознанной не является.

Насколько легко должна быть извлекаема информация, чтобы попасть под определение автора? Это можно показать на следующем примере. Предположим, доктор Ноу собирается убить Джеймса Бонда и для этого конструирует сложную машину. В процессе создания машины он сильно увлекается работой, так что не думает об убийстве. Тем не менее, если бы в этот момент кто-то спросил его о том, чем он занимается, он бы ответил: «Я пытаюсь убить Джеймса Бонда». По мнению Леви, эта информация является персонально доступной доктору Ноу: даже не обдумывая ее в данный момент, он легко может ее извлечь. Данного типа доступ к такого рода информации необходим для моральной ответственности. «Агент должен осознавать моральное значение собственных поступков»<sup>8</sup>, — уточняет свой «тезис о сознании» автор. В таком случае в чем спорность этого тезиса, если выводы Вегнера и Либета ему не противоречат? Кто же опонирует Леви?

Основными противниками тезиса о сознании являются не Вегнер и Либет, а коллеги-философы. Это сторонники так называемой экспрессивистской позиции в отношении моральной ответственности Номи Арпэли, Анжела Смит, Джордж Шер<sup>9</sup>:

8. Ibid. P. 37.

9. См.: Sher G. Who Knew? Responsibility without Awareness. Oxford: Oxford University Press, 2009.

Такие философы, как Арпэли, Смит, Шер и другие оппоненты тезиса о сознании, отрицают, что информация, которая влияет на управление поведением, должна быть легко и непосредственно доступна агенту, чтобы он был морально ответствен за это поведение<sup>10</sup>.

Экспрессивисты, вероятно, сочли бы не только доктора Ноу ответственным за попытку убийства, но и Эмили — за игнорирование годовщины свадьбы. Арпэли в книге «Добродетель без принципов: исследование морального агента»<sup>11</sup> приводит сходный пример с Гекльберри Финном.

Финн считает своим моральным долгом передать беглого раба Джима властям, поскольку Джим — чья-то собственность, а воровать аморально. Однако Финн не может найти в себе силы сделать то, что считает морально оправданным, и вместо этого содействует побегу Джима. Финн не может извлечь информацию, которая мотивирует его поступок, однако заслуживает за него благодарность. Последняя, следовательно, не требует, чтобы моральные мотивы были доступны агенту, а значит, это не необходимо и для моральной ответственности в целом, считает Арпэли. Леви с этим выводом не согласен. Но как разрешить спор? Интуиции в данном случае действительно могут различаться. Аргумент Леви построен на представлении о сознании как глобальном рабочем пространстве.

Нейрофизиологические исследования показывают, что мозг не является единым механизмом для обработки информации, а состоит из множества дискретных частей. Нарушение в одном модуле порождает лишь локальные дефекты, но не прекращение работы системы в целом. Обработка информации в модулях происходит не последовательно, а параллельно, и нет одного физического места, куда вся информация стекалась бы «для представления агенту». Такая физическая архитектура вполне оправдана с эволюционной точки зрения: она обеспечивает скорость реакции и восприятия, а также надежность когнитивной системы. Но она плохо соотносится с общими принципами рационального мышления, работающего вне зависимости от специфических зон экспертизы. «Организация такого типа плохо подходит для обработки информации вне определенной области»<sup>12</sup>. Рациональное

10. *Levy N. Consciousness and Moral Responsibility.* P. 35.

11. См.: *Arpaly N. Unprincipled Virtue. An Inquiry into Moral Agency.* N.Y.: Oxford University Press, 2004.

12. *Levy N. Consciousness and Moral Responsibility.* P. 41.

мышление также предполагает, что, оценивая какое-либо утверждение, мы сопоставляем его с другими имеющимися у нас представлениями. Но как все это возможно в распределенной модульной системе? Как может единый разум быть супервентным на модульной дискретной архитектуре?

На эти вопросы позволяет ответить модель глобального рабочего пространства, предложенная когнитивистом Барнардом Барсом и доработанная впоследствии в нейробиологической теории Дехана и Накачче. Согласно этой теории, осознание представляет собой процесс трансляции информации большому числу разнообразных модулей.

По мнению Дехана и его коллег, эти парадигмы показывают, что сознание связано с глобальным нейронным пространством, которое образовано в мозге масштабной сетью расположенных в разных зонах кортикальных нейронов, передающих импульсы другим гомологичным нейронам через длинные возбужденные аксоны<sup>13</sup>.

Глобальное рабочее пространство возникает как некоторая единая виртуальная сеть на базе нейронной модульной архитектуры мозга. Но что дает интеграция информации с помощью виртуальной сети? По мнению Леви, интеграция позволяет, во-первых, формировать более гибкую реакцию на внешние стимулы. И во-вторых, она обеспечивает такую реакцию, которая соответствует общим представлениям агента, а не только алгоритмам работы одного или нескольких его модулей.

Бессознательное поведение людей (например, в ситуации отсутствия внимания или сомнамбулизма) управляется сценариями действий. Это заученные моторные автоматизмы, которые освобождают от излишней вычислительной работы, например, при ходьбе, беге, управлении автомобилем (переключении передач, сигнализации о повороте и т. п.), питании, использовании различных орудий, игре на музыкальных инструментах и т. д. Существуют сложные комплексные сценарии действий, но и они большей частью носят баллистический характер: после запуска они невосприимчивы к принципиально новым внешним стимулам.

Подлинная гибкость реакции — чувствительность к содержанию разнообразных сигналов... требует сознания<sup>14</sup>.

13. Ibid. P. 49.

14. Ibid. P. 76.

Ресурсы одного модуля, доступные без участия сознания, неадекватны для обработки нестандартной ситуации и формирования соответствующей реакции — для этого необходимо глобальное рабочее пространство. Задействование последним большого набора модулей расширяет диапазон доступных реакций. Сознание как глобальное рабочее пространство необходимо для обработки информации «на персональном уровне». В результате интеграции информации

... действия, управляемые осознанными представлениями, с большей вероятностью отражают общность ментального содержания агента<sup>15</sup>.

Это обеспечивает большую гибкость не только моторных реакций, но и мышления, что позволяет агенту приблизить свое поведение к оптимальному рациональному идеалу, учесть различные внешние обстоятельства и согласовать реакцию со множеством внутренних представлений, ценностей и установок. Таким образом, сознание обеспечивает возможность более адекватной поведенческой реакции, отвечающей установкам агента в целом.

Сознание необходимо для того, чтобы поведение могло быть более чувствительным... к нашим установкам на личностном уровне<sup>16</sup>.

Описанные функции сознания необходимы для приписывания агенту моральной ответственности, что следует из анализа двух основных теорий моральной ответственности — «теории самовыражения реального „Я“» и «теории контроля». Сторонники первой считают, что ответственными являются только такие действия, которые выражают моральные оценки агента (*evaluative agency*). Моральные оценки направляют поведение агента, поэтому с их помощью можно относительно надежно его предсказывать. Поведение, продиктованное моральными оценками и суждениями, выражает моральную природу субъекта, но лишь сознание позволяет таким оценкам влиять на поведение. Для выражения моральных суждений личности необходимо, чтобы агент был в сознании (*creature consciousness*) и осознавал моральные аспекты собственных поступков (*state consciousness*). Действия, совершен-

15. Ibid. P. 83.

16. Ibid. P. 86.

ные без соблюдения этих условий, управляются только локальными сценариями и бессознательными установками. Ни первые, ни вторые не чувствительны к разумным основаниям: они возникают и изменяются, минуя рациональный контроль. Следовательно, они не могут выражать моральные суждения и оценки агентов. То есть поступки, совершенные под влиянием сценариев действий и бессознательных установок, не являются ответственными. Таким образом, «тезис о сознании» должен быть верным для сторонников теории моральной ответственности, основанной на «самовыражении реального „Я“».

Сторонники второй теории моральной ответственности считают, что ответственность могут влечь только те действия, которые агент контролирует в достаточной степени. Под достаточным контролем исторически чаще всего понимали возможность поступить иначе, однако сегодня этому фактору уже не придают большого значения. Наиболее влиятельный проponent современного взгляда на эту проблему философ Джон Фишер под необходимым контролем предлагает понимать возможность «направляющего контроля» (*guidance control*) действиями — восприимчивость к доводам, в том числе моральным, как основаниям поступать иначе. Агент управляет действием тогда, когда в контрафактической ситуации поступил бы иначе, имея на то соответствующие основания:

Мы управляем действиями тогда, когда чувствительны к разнообразным соображениям в отношении этих действий, включая моральные, и воспринимаем их как основание поступить иначе; и когда действительно поступили бы иначе, будь у нас другие соображения в контрафактической ситуации<sup>17</sup>.

Фишер и другие сторонники такого подхода к теории моральной ответственности не настаивают на том, чтобы действия выражали моральные оценки агента. Но и они, по мнению Леви, требуют, чтобы во время совершения действия агент пребывал в сознании и, более того, понимал моральные аспекты своих поступков. Во-первых, очевидно, что без сознания агент невосприимчив к большинству рациональных доводов. Во-вторых, без осознания моральных аспектов действий агент не может увидеть расхождения между моральными доводами и решениями о действии.

17. Ibid. P. 83.

Когда наше поведение диктуется неосознанными факторами, мы не можем быть к ним чувствительны и обнаруживать противоречия или согласие между этими и другими мотивами<sup>18</sup>.

Таким образом, согласно второй теории моральной ответственности, мы управляем поведением, только когда находимся в сознании, и управляем только теми действиями, моральные основания которых осознаем. Следовательно, и в этом случае «тезис о сознании» находит свое подтверждение. Какими же будут логические следствия из принятия тезиса Нила Леви?

«Тезис о сознании» может иметь следствия для принятия юридических решений и моральных суждений о виновности индивида. В частности, если позиция Леви справедлива, то преступления, совершенные без состояния осознанности, не должны влечь наказание. К таким преступлениям относится, например, известный случай с Кеннетом Парксом и ему подобные. Случай Паркса особенно примечателен. Летом 1987 года Паркс встал рано утром, проехал 23 км на автомобиле, ворвался в дом родителей своей жены, убил тещу и почти задушил тестя. Как показало следствие, все это Паркс совершил в состоянии сомнамбулизма, которым страдал он сам и его родственники. Паркс не имел намерения совершать преступление, он очень любил родителей жены и самостоятельно приехал в полицию, где заявил, что «кажется, кого-то убил». Согласно «тезису о сознании», действия Паркса были совершены под управлением сценариев действия, не могли выражать его моральные установки и не были управляемы им. Следовательно, эти действия не подпадают под разряд ответственных.

Следствием «тезиса о сознании» должно стать освобождение от прямой ответственности людей, совершающих и более тривиальные поступки, например выносящих суждения под действием бессознательных установок (даже если они выражают скрытый сексизм, расизм, отсутствие толерантности) или действующих по забывчивости (как забывшая о годовщине свадьбы Эмили). В подобных ситуациях агенты не способны анализировать соответствие установок другим убеждениям, и потому их действия не могут быть ответственными. Практическое значение и ценность «тезиса о сознании» и исследования Леви очевидны. Впрочем, нельзя не обратить внимания и на некоторые потенциально уязвимые места его теории. Возражения могут исходить от двух

18. Ibid. P. 115.

господствующих, но взаимоисключающих подходов к теории сознания — от квалюфилов и иллюзионистов.

Часть философов, работающих над проблемой сознания, в частности Джон Сёрл, Дэвид Чалмерс, Томас Нагель и др., считают, что сознание — это субъективные квалитативные состояния. Странники этой позиции, которых Дэниел Деннет назвал квалюфилами, как правило, полагают, что ментальные, интенциональные содержания существуют только тогда, когда есть квалиа. Согласно этой точке зрения, одних лишь информационных процессов, осведомленности агента или способности к вербальным отчетам недостаточно для обеспечения семантики. Но в таком случае глобальное рабочее пространство не выполняет функции согласования семантического контента. Из теории Леви следует, что суждения об ответственности обычного агента и функционального изоморфа, философского зомби, должны быть идентичными, с чем квалюфилы, скорее всего, не согласились бы. Они, вероятно, заявили бы, что попадание сигналов в глобальное рабочее пространство не дает принципиального различия и не представляет оснований для согласования убеждений. Интеграция убеждений на персональном уровне возможна только при наличии феноменального опыта.

Впрочем, слабые места в позиции Леви могут найти и противники квалюфилов — иллюзионисты во главе с Деннетом, считающие квалиа иллюзией. Они могли бы не согласиться с выводами Леви потому, что иначе идентифицируют контент, ментальные семантические содержания. Так, по определению Деннета, убеждения можно приписывать в случае, если они должны существовать у организма, учитывая место агента в мире и его интересы, а также если эти приписанные убеждения позволяют надежно предсказывать поведение агента<sup>19</sup>. Но тогда, видимо, под понятие убеждений агента подпадают и «бессознательные установки» Леви. Они помогают надежно предсказывать поведение агента в определенных ситуациях и должны были бы существовать, учитывая интересы и место агента. Но это означает, что «бессознательные установки» выражают в том числе моральные оценки агента и должны быть чувствительны к другим его мотивам. Следовательно, действия под управлением бессознательных установок подпадают под тип ответственных, а «тезис о сознании» ложен.

Эти возможные линии критики или расхождения с другими теоретиками сознания едва ли умаляют достоинства работ

19. См.: *Dennett D. Intentional stance. Cambridge, MA: MIT Press, 1989.*

Нила Леви, которые остаются важнейшими исследованиями для всей проблематики свободы воли, моральной ответственности и сознания.

### *Библиография*

- Arpaly N. Unprincipled Virtue. An Inquiry into Moral Agency. N.Y.: Oxford University Press, 2004.
- Dennett D. Intentional Stance. Cambridge, MA: MIT Press, 1989.
- Levy N. Consciousness and Moral Responsibility. Oxford: Oxford University Press, 2014.
- Levy N. Hard Luck: How Luck Undermines Free Will and Moral Responsibility. Oxford: Oxford University Press, 2015.
- Levy N. Neuroethics: Challenges for the 21st Century. Cambridge, UK: Cambridge University Press, 2007.
- Libet B. Mind Time: Temporal Factors of Consciousness. Cambridge, MA: Harvard University Press, 2005.
- Sher G. Who Knew? Responsibility without Awareness. Oxford: Oxford University Press, 2009.
- Wegner D. The Illusions of Conscious Will. Cambridge, MA: MIT Press, 2003.

## THE “CONSCIOUSNESS THESIS” AND MORAL RESPONSIBILITY IN NEIL LEVY’S RESEARCH

DMITRY VOLKOV. Co-director, Moscow Center for Consciousness Studies.  
Address: 14A Burdenko str., 119121 Moscow, Russia. E-mail: dvolkoff@gmail.com.

*Keywords:* free will; moral responsibility; consciousness; consciousness thesis; Neil Levy.

This paper presents an analysis of the “consciousness thesis” by Neil Levy and the argument in support of it. The “consciousness thesis” is a claim that consciousness is a necessary condition for moral responsibility. The author of the argument uses the word “consciousness” to refer to the information process in the brain that affects the behavior of the agent, and the contents of which can be reported verbally by him or her. The opponents of this thesis include philosophers-expressivists like Nomi Arpaly, Angela Smith, George Sher. Levy’s argument is built on the Global Workspace Model of consciousness. Levy believes that only consciousness as a global workspace can provide integration of information on the personal level. Automatism and behavior not guided by conscious understanding of moral aspects of actions are not integrated on the personal level. This integration, however, supports the flexibility and which is necessary for moral responsibility according to two dominant theories: Real Self theory and control theory. Therefore, the consciousness thesis is proved.

The author of this paper suggests that the weak point of Levy’s argument is his definition of consciousness. The consciousness thesis is most likely to be refuted by the proponents of two main theories of consciousness: qualophiles and illusionists. Qualophiles claim that the semantic content is based on the qualitatively phenomenal states, and integration of information by itself is not enough to form beliefs on the personal level. Therefore, consciousness as defined by Levy is not relevant for forming personal-level beliefs. Therefore it is not necessary for moral responsibility. Illusionists claim that beliefs are ascribed to the agents from the outside when their behavior is being interpreted. So even those processes that cannot easily be verbally reported still represent personal beliefs. From this we can conclude that consciousness as defined by Levy is not a necessary condition for moral responsibility.

### References

- Arpaly N. *Unprincipled Virtue. An Inquiry into Moral Agency*, New York, Oxford University Press, 2004.
- Dennett D. *Intentional Stance*, Cambridge, MA, MIT Press, 1989.
- Levy N. *Consciousness and Moral Responsibility*, Oxford, Oxford University Press, 2014.
- Levy N. *Hard Luck: How Luck Undermines Free Will and Moral Responsibility*, Oxford, Oxford University Press, 2015.
- Levy N. *Neuroethics: Challenges for the 21st Century*, Cambridge, UK, Cambridge University Press, 2007.
- Libet B. *Mind Time: Temporal Factors of Consciousness*, Cambridge, MA, Harvard University Press, 2005.
- Sher G. *Who Knew? Responsibility without Awareness*, Oxford, Oxford University Press, 2009.
- Wegman D. *The Illusions of Conscious Will*, Cambridge, MA, MIT Press, 2003.