

# ИНТЕЛЛЕКТУАЛЬНЫЕ ПОРОКИ КАК НЕЯВНЫЕ УСТАНОВКИ\*

**Беседин Артем Петрович** – кандидат философских наук, доцент.  
Московский государственный университет имени М.В. Ломоносова, Российская Федерация, 119234, г. Москва, Ломоносовский просп., д. 27/4;  
e-mail: [a.besedin@hardproblem.ru](mailto:a.besedin@hardproblem.ru)



В статье анализируется важное для современной эпистемологии добродетелей понятие интеллектуального порока – черты интеллектуального характера, препятствующей ответственному исследованию. Цель данной статьи – сформулировать гипотезу о том, что в настоящее время в современной культуре значительная часть эпистемических пороков представляет собой неявные установки. В первой части статьи исследуется понятие неявной установки, разбираются примеры неявных установок (неявный сексизм и расизм). Во второй части статьи показывается, что 1) в случаях неявно предвзятого поведения имеет место проявление эпистемического порока; 2) «мотивационная» теория, опирающаяся на идеи Л. Загзебски, не может объяснить проявление интеллектуального порока в случаях неявной предвзятости. В третьем разделе статьи продемонстрировано, что неявные установки могут быть чертами характера (подобно моральным порокам, не признаваемым самим субъектом) – приобретенными, укорененными в личности, но все же корректируемыми. Четвертая часть статьи посвящена разбору того, при каких условиях интеллектуальные пороки могут быть явными установками. Делается вывод о том, что это возможно, если агент является «дьявольским существом» (руководствуется злом как целью), имеет эгоистические пороки (применяет разные критерии порочности к себе и к другим) или иррационален. В заключении показывается, что распространение критического мышления в современном обществе должно вести к переходу интеллектуальных пороков из явных установок в неявные.

**Ключевые слова:** эпистемология добродетелей, эпистемология пороков, интеллектуальные пороки, неявные установки, неявная предвзятость

## INTELLECTUAL VICES AS IMPLICIT ATTITUDES

**Artem P. Besedin** – PhD in Philosophy, Associate Professor.  
Lomonosov Moscow State University,  
27-4 Lomonosovsky Avenue  
119234, Moscow,  
Russian Federation;  
e-mail: [a.besedin@hardproblem.ru](mailto:a.besedin@hardproblem.ru)

The article analyzes an important concept of contemporary virtue epistemology – the concept of intellectual vice, that is a trait of intellectual character that hinders responsible research. The purpose of this article is to formulate a hypothesis that, today, in the modern culture, a significant part of epistemic vices are implicit attitudes. The first part of the article explores the concept of implicit attitude, examines examples of implicit attitudes that have become widespread in the research literature: implicit sexism and racism. The second part of the article shows that in cases of implicitly biased behavior there is a manifestation of epistemic vice, and that the “motivational” theory based on Zagzebski’s

\* Исследование выполнено за счет гранта Российского научного фонда № 21-78-10044 «Феномен моральной ответственности».



ideas cannot explain the manifestation of intellectual vice in cases of implicit bias. In the third section of the article, it is demonstrated that implicit attitudes can be traits of the agent's character (like moral vices that are not recognized by the subject herself): they can be acquired, rooted in the personality, and can be corrected. The fourth paragraph of the article analyses the conditions under which intellectual vices can be explicit attitudes. It is possible if the agent is a diabolical being (guided by evil as a goal), has egoistic vices (applies different criteria of vice to himself and to others), or is irrational. In the final section, it is shown that the spread of critical thinking in modern society should lead to the transition of intellectual vices from explicit attitudes to implicit ones. The conclusion is made about the theoretical and practical significance of the hypothesis under discussion. From a theoretical point of view, it allows us to explain why intellectual vices are widespread and difficult to eradicate, to place vices as character traits between the local (situational) and global levels, to apply to the concept of epistemic vice all the research concerning implicit attitudes, to develop a theory of epistemic responsibility. In practical terms, this hypothesis can be used to analyze the manifestations of implicit vices in various spheres.

**Keywords:** virtue epistemology, vice epistemology, intellectual vices, implicit attitudes, implicit bias

Понятие интеллектуального (эпистемического) порока является столь же важным для современной эпистемологии добродетелей, как и понятие самой интеллектуальной добродетели. Кассим Кассам в своей влиятельной статье «Эпистемология пороков» дает им такое определение: «Пороки интеллектуального характера – это черты интеллектуального характера, которые препятствуют эффективному и ответственному исследованию» [Cassam, 2016, p. 164]. Эффективным он называет такое исследование, которому присуща *знаниепроводимость* (*knowledge-conduciveness*) [Ibid., p. 166]. Ответственное исследование «руководствуется свидетельствами и признает обязательства, свойственные исследователю» [Ibid., p. 164]. Изучение интеллектуальных пороков является особенно важным, потому что пороки, по-видимому, распространены среди людей гораздо шире, чем добродетели. Примеры интеллектуальных пороков столь же многочисленны, как и примеры их моральных аналогов. Линда Загзевски перечисляет такие: «Интеллектуальная гордыня, пренебрежение, праздность, трусость, конформизм, беззаботность, черствость, предубеждение, предвзятое мышление, закрытость ума, нечувствительность к деталям, тупость и отсутствие аккуратности» [Zagzebski, 1996, p. 152]. Кассам добавляет пороки догматизма и доверчивости [Cassam, 2016, p. 159].

Как показывают эти примеры, понятие эпистемического порока является довольно широким и требует уточнения. В представленной форме оно уязвимо для критики со стороны эпистемического ситуационизма



и антиглобалистского<sup>1</sup> подхода к характеру. Вызов со стороны ситуационизма состоит в том, что определенные черты могут проявляться у человека, когда он находится в подходящей ситуации и именно из-за нахождения в этой ситуации, а не потому, что эта черта присуща его характеру. Антиглобалистская критика опирается на тезис, что мы почти никогда не обладаем совершенно неизменными чертами характера, проявляющимися одинаково в одинаковых ситуациях. Всё это ставит под сомнение объяснительный потенциал такого понятия, как «черта характера», а вместе с ним и понятий интеллектуальных пороков и добродетелей. Кассам предлагает свои варианты ответа на эти вызовы и подводит такое резюме: «Золотая середина находится где-то между мифическими глобальными чертами, которые требуют абсолютной согласованности, и ультратонкими, ситуативно-определенными локальными чертами, которые ничего не говорят о поведении человека, кроме как в одном-единственном случае» [Cassam, 2016, p. 174].

В этой статье я буду защищать очень слабый тезис: в настоящее время в современной культуре значительная часть эпистемических пороков представляет собой неявные установки. Как таковой этот тезис может рассматриваться как эмпирически проверяемая гипотеза. Статья начинается с рассмотрения феномена неявных установок и случаев неявно предвзятого поведения. Я приведу аргументы в пользу того, что во многих случаях эпистемические пороки являются неявными установками. Такое их понимание дает нам возможность разместить многие интеллектуальные пороки между уровнем локальных проявлений и глобальных черт характера, выполнив требование «золотой середины» Кассама. Будет проведено различие между явными и неявными интеллектуальными пороками. В заключении я постараюсь показать общую продуктивность трактовки эпистемических пороков как неявных установок.

## Неявные установки и неявная предвзятость

Наши действия являются проявлениями наших установок. Понятие «установки» широко используется в современной философии и психологии. Психолог Дэвид Майерс описывает его так: «Когда люди спрашивают о чьих-то установках, они имеют в виду убеждения и чувства, возникающие в связи с каким-нибудь человеком или событием, и, как следствие, готовность вести себя определенным образом. Взятые вместе, благоприятные или неблагоприятные оценочные реакции на что-либо, независимо от того, в какой форме они выражены (в форме убеждений, чувств или готовности к действиям), определяют

---

<sup>1</sup> Эти направления критики объединяются в работе Марка Альфано [Alfano, 2012].



установку человека» [Майерс, 2010, с. 160]. Система установок деятеля двойственна: «Относительно одного и того же предмета у нас могут быть явные (осознанные) и неявные (автоматические) установки» [Там же, с. 393]. Философ Нил Леви характеризует неявные установки как такие, которые «активируются автоматически, не могут быть намеренно подавлены и, при этом, не признаются агентом открыто» [Levy, 2015, p. 817]. Поведение, в котором проявляются неявные установки, я буду называть неявно предвзятым поведением.

Сегодня наибольшее внимание психологов и философов привлекают неявные расовые и гендерные стереотипы. Открытое проявление предвзятого отношения к индивиду на основании того, что он или она принадлежит к некоторой социальной группе, в современных западных обществах уже много десятилетий считается неприемлемым. Однако исследования в области социальной психологии показывают, что те установки, которые около века назад многие открыто принимали, сегодня неявно влияют на поведение людей, искренне считающих себя приверженцами противоположных убеждений. За последние десятилетия было проведено множество экспериментов, показывающих, что люди, не являющиеся явными сексистами или расистами, в определенных смоделированных в рамках эксперимента условиях склонны к проявлению неявной дискриминации по гендерному или расовому признаку.

Среди таких экспериментов особое распространение получили так называемые *исследования резюме (CV studies)*. Например, в ходе эксперимента Эрика Ульмана и Джеффри Коэна [Uhlmann & Cohen, 2005] участникам было предложено оценить резюме кандидатов при приеме на работу на стереотипные мужские или женские должности. Участники, независимо от пола, были склонны отдавать предпочтение мужчинам при подборе кандидатов на стереотипно мужскую работу и женщинам – на стереотипно женскую. Далее организаторы эксперимента опрашивали участников на предмет того, считают ли они свой выбор объективным. Большинство испытуемых демонстрировало высокую степень уверенности в объективности своего выбора, а когда им задавали вопросы о критериях выбора, они «подстраивали свои критерии так, чтобы придать больший вес качествам конкретного аппликанта того пола, которому отдается предпочтение, независимо от того, являются ли эти характеристики стереотипно мужскими или женскими» [Ibid., p. 479]. Один из выводов Ульмана и Коэна состоит в том, что испытуемые находились под действием «иллюзии объективности» [Ibid., p. 477] и не были осведомлены о проявлении неявной установки.

Майкл Нортон, Джозеф Ванделло и Джон Дарли провели серию экспериментов, аналогичных эксперименту Ульмана и Коэна, но ориентированных на изучение расовых стереотипов [Norton et al., 2004]. То, что Ульман и Коэн назвали «подстраиванием критериев», Нортон



с соавторами обозначили термином «казуистика». В ходе этого исследования было показано, что испытуемые, давая казуистические объяснения выбору предпочтительного кандидата, часто приводили искаженную информацию о кандидатах. Кроме того, Нортон и соавторы продемонстрировали зависимость между склонностью к предвзятому выбору и эластичностью ситуации выбора [Norton et al., p. 819–822]. Если между двумя кандидатами с разным цветом кожи имеется одно существенное различие (например, в уровне образования), а все остальные характеристики совпадают, то выбор между этими кандидатами неэластичен. Испытуемые, сталкиваясь с ситуацией неэластичного выбора, были склонны отдавать предпочтение лучшему кандидату независимо от цвета кожи. Но в реальных ситуациях выбор часто эластичен: это означает, что между кандидатами имеется множество различий. Будучи поставлены в эластичные условия, испытуемые часто использовали не имеющие отношения к делу различия между кандидатами для объяснения своего предвзятого решения.

Другие эксперименты основаны на различных вариациях теста на неявные ассоциации (см. [Brownstein, 2019, web, §1.2]). Исследования Кита Пейна касаются феномена *предубежденности относительно оружия* (результаты экспериментов суммированы в [Payne, 2006]). В частности, было показано, что в экспериментальных условиях испытуемые (независимо от собственного цвета кожи) склонны расценивать неопределенный предмет как оружие скорее в руках чернокожего, чем в руках белого. Важным результатом исследований Пейна и коллег является вывод о том, что «предвзятость может существовать с осознанным намерением быть справедливым и непредвзятым» [Ibid., p. 288]: даже если испытуемым перед экспериментом говорили, чтобы они избегали предвзятых действий, они продолжали их совершать. Случаи неявно предвзятого поведения по отношению к представителям другого пола или расы лучше всего исследованы в литературе, однако неявно предвзятое поведение может иметь место в отношении инвалидов, престарелых людей, ЛГБТ и представителей других социальных групп.

Наиболее существенные черты неявно предвзятого поведения были отмечены выше в цитате из статьи Леви: автоматизм, неосознанность, отсутствие у агента контроля над ними. Это отличает неявную установку от явной, которую можно охарактеризовать как «относительно постоянный и общий оценочный ответ, о котором личность сознательно осведомлена» [American Psychological Association, web]. Однако, как отмечают де Хувер и соавторы, понятия «автоматического» и «неявного» не нужно отождествлять: «Например, есть подтверждения тому, что активация стереотипов происходит автоматически в том смысле, что она не зависит от осознаваемой цели активировать стереотип или от наличия ресурсов для ее обработки,



но она не автоматическая в том отношении, что она зависит от наличия некоторых других целей» [De Houwer et al., 2009, p. 350]. Важной чертой явных установок является то, что субъект склонен открыто отражать в своих утверждениях содержание этих установок [Levy, 2012, p. 31]. «Явные установки, – также пишет Леви, – в тех случаях, когда они используются в качестве посылок в рассуждениях, требующих усилий, производят контент способом, подчиняющимся правилам, которые агент (имплицитно) принимает в качестве нормативных» [Ibid., p. 32].

В дальнейшем я буду исходить из допущения, что неявные установки не являются убеждениями в полном смысле этого слова. «Большинство современных философов, – пишет Эрик Швитцгебель, – характеризуют убеждение как *пропозициональную установку*... Пропозициональная установка – это ментальное состояние, связанное с наличием установки, предрасположенности, принятия, мнения о пропозиции или о потенциальном положении дел, в котором эта пропозиция будет истинной» [Schwitzgebel, 2019]. Явные установки агента – это его убеждения, то есть это пропозициональные установки, выражаемые в готовности агента соглашаться с одними суждениями и не соглашаться с другими, защищать одни и оспаривать другие. Неявные установки – это не убеждения в этом смысле.

Швитцгебель характеризует неявные установки как промежуточные убеждения (*in-between beliefs*) [Schwitzgebel, 2010], Леви выделяет их в отдельную категорию когнитивных состояний, для которой предлагает название «обрывочные предпочтения» (*patchy endorsements*) [Levy, 2015], Тамар Гендлер предлагает считать их праубеждениями (*aliefs*) [Gendler, 2008]. Концептуализация неявных установок Гендлер представляется наиболее плодотворной для целей данной статьи.

Согласно Гендлер, праубеждение – это ментальное состояние, главными характеристиками которого являются ассоциативность, автоматизм (субъект может быть сознательно осведомлен о своем праубеждении, но праубеждения работают без вмешательства сознания) и арациональность (*arationality*) (праубеждения могут быть полезными или вредными, похвальными или заслуживающими презрения, но они не рациональны и не иррациональны). Праубеждения предшествуют другим когнитивным установкам и часто аффективно нагружены [Ibid., p. 557–558].

Гендлер противопоставляет праубеждения убеждениям. Она не дает точного определения убеждения, лишь характеризуя его следующим образом: «В случае изменения имеющихся данных, учитывающих все обстоятельства, убеждение подвергается немедленному пересмотру» [Ibid., p. 565]. Праубеждения не таковы: они не меняются сразу же при изменении имеющихся у субъекта данных. Тем не менее праубеждения в принципе могут изменяться. Гендлер выделяет два



способа корректировки праубеждений: аристотелевский и картезианский. Первый предполагает формирование противоположных привычек путем намеренного повторения; второй – перенос внимания при помощи создания в воображении образа желаемого результата [Gendler, 2008, p. 554]. Праубеждения могут соответствовать убеждениями агента, а могут не соответствовать [Ibid., p. 570–572]. В первом случае мы имеем случай гармонии между явными и неявными установками субъекта.

Из всех приведенных выше примеров неявно предвзятого поведения для дальнейшего анализа я выберу сценарии исследований резюме и буду опираться на результаты, полученные Ульманом, Коэном и исследовательской группой Нортон. В этих сценариях у испытуемых было достаточно времени для обдумывания и принятия решения, они должны были выполнить довольно рутинную задачу в привычных условиях. Это отличает условия исследований резюме от исследований предубежденности относительно оружия, когда участникам нужно быстро принять решение, от которого может зависеть их жизнь. В последнем случае можно усомниться в том, что у испытуемых в принципе есть возможность провести ответственное исследование. Исследования резюме, наоборот, предполагают проведение такого исследования как часть парадигмы эксперимента. Поэтому именно сценарии исследований резюме представляют наибольший интерес в эпистемологическом контексте.

## **Имеет ли место проявление интеллектуального порока в случаях неявно предвзятого поведения?**

Случаи дискриминации, очень похожие на те, что я привел выше, разбираются Мирандой Фрикер в ее книге «Эпистемическая несправедливость» [Fricker, 2007]. Прежде чем перейти к дальнейшему обсуждению, я должен особо подчеркнуть, что примеры неявной предвзятости, хотя они, очевидно, похожи на сценарии открытой дискриминации, рассматриваются мной не как случаи эпистемической несправедливости<sup>2</sup>, а как проявления обычных эпистемических пороков. Во-первых, эпистемическая несправедливость сама по себе не является интеллектуальным пороком, поскольку это свойство практик, принятых в сообществе. Во-вторых, случаи неявной предвзятости не подходят под узкие рамки исключительно эпистемической несправедливости, заданные Фрикер. Такая несправедливость

<sup>2</sup> Об эпистемической несправедливости см. [Шевченко, 2020], а также дискуссию по этой статье в том же номере «Эпистемологии и философии науки».



«в наиболее фундаментальном смысле заключается в чем-то неправильном, совершенном в отношении кого-то как обладающего способностью к познанию» [Fricker, 2007, p. 1]. Неявная предвзятость может вести и часто ведет к эпистемической несправедливости в этом смысле, однако используемые в этой статье примеры показывают, что первая, очевидно, не является видом второй.

Если следовать приведенному в начале статьи определению эпистемического порока как черты интеллектуального характера, препятствующей эффективному и ответственному исследованию, то можно заметить, что в случаях неявно предвзятого поведения мы сталкиваемся именно с интеллектуальными пороками. Действительно, неявная предвзятость – это такое свойство агента, которое не дает ему или ей осуществлять знаниепроедливое исследование (*knowledge-conducive inquiry*).

В пользу того вывода, что описываемые случаи включают проявления эпистемических пороков, говорит и та терминология, которую предпочитают использовать авторы процитированных исследований. Так, Ульман и Коэн пишут об *иллюзии объективности*, а Нортон и соавторы – о *казуистике* со стороны испытуемых. Этим подчеркивается то, что под влиянием неявной предвзятости агенты подгоняли обоснование под свое решение, то есть действовали эпистемически порочно.

Если неявная предвзятость может быть интеллектуальным пороком, то перестает работать способ определения эпистемических пороков через структуру мотивации агента, которого придерживаются, например, Мейсон и Уилсон, опираясь на Загзебски. Загзебски утверждает, что интеллектуальные пороки связаны с недостаточной мотивацией к знанию [Zagzebski, 1996, p. 170]. Мейсон и Уилсон развивают эту идею и определяют интеллектуальный порок по аналогии с моральным: «Моральные добродетели – это те черты характера, которые необходимо связаны с глубокой мотивацией к морально значимым целям, а моральные пороки – это те черты, которые необходимо связаны с провалом в мотивации такого рода. Эпистемические добродетели – это те черты характера, которые связаны с глубокой мотивацией к эпистемически значимым целям; эпистемические пороки – это те, которые связаны с провалом мотивации такого рода» [Mason & Wilson, 2017, p. 88]. Говоря об эпистемически значимых целях, Мейсон и Уилсон цитируют Загзебски [Zagzebski, 1996, p. 166–194] и указывают на *когнитивный контакт с реальностью*, подразумевающий такие эпистемические ценности, как *истинное убеждение, знание и понимание* [Mason & Wilson, 2017, p. 88]. Мейсон и Уилсон специально поясняют, что они имеют в виду под *глубокой мотивацией*: во-первых, она должна быть *постоянной*, во-вторых, она должна *сильно ощущаться*, в-третьих, она должна быть *устойчивой* [Ibid., p. 87].





Мотивационная теория (назовем ее так) не редуцирует добродетели и пороки к структуре мотивации агента (теория Загзебски кроме мотивационного компонента включает компонент успешности интеллектуальных добродетелей [Zagzebski, 1996, p. 176–194]). Такая теория, как я ее определяю, утверждает лишь то, что между интеллектуальными пороками и добродетелями есть существенное различие в структуре мотивации, и именно это различие делает порок пороком, а добродетель – добродетелью. Еще одна важная оговорка: сторонник мотивационной теории не обязан утверждать, что эпистемический порок – это всего лишь противоположность добродетели. Загзебски это прямо отрицает: «Порок – это нечто контрарное добродетели, а не контрадикторное... [П]орок – это не просто отсутствие добродетели. Это так, потому что порок, как и добродетель, приобретается через привыкание» [Ibid., p. 112]. Наконец, я не оспариваю того, что мотивационная теория может успешно объяснять природу интеллектуальных добродетелей, однако прямое ее применение к порокам, по-видимому, дает сбой.

Мотивационная теория не может включить неявную предвзятость в класс интеллектуальных пороков. В экспериментах Ульмана и Коэна, Нортона и его коллег испытуемые имели достаточно глубокую мотивацию провести ответственное исследование и выбрать лучшего кандидата. Это предполагается парадигмой эксперимента: авторы экспериментальных исследований исходили из того, что испытуемые имели достаточно устойчивую мотивацию к эпистемически значимым целям (никто из них не пытался намеренно обмануть исследователей) и были сконцентрированы на выполнении задания (то есть ощущали сильную мотивацию выбрать лучшего кандидата и не переключались на другие цели из-за дуновения ветерка). Допущение обратного означало бы, что в поставленных экспериментах участвовали плохо мотивированные, невнимательные испытуемые, и их было настолько много, что их результаты существенно повлияли на общую статистику. Разумеется, это абсолютно необоснованное допущение. Кроме того, эпистемический порок, который проявляли испытуемые в экспериментальных ситуациях, – это вовсе не тот порок, который Мейсон и Уильямс приписывают носителям культурных стереотипов: это не эпистемическая лень или нежелание выходить из «зоны комфорта»<sup>3</sup>. Перед испытуемыми в обсуждаемых экспериментах были поставлены довольно простые задачи, не предполагающие серьезных вложений в исследование.

<sup>3</sup> «Многие из наших “когнитивных” ошибок относительно того, как устроен мир, могут быть поняты как пороки лени – то есть тенденции к тому, чтобы позволить нашему стремлению к интеллектуальному комфорту подавлять наше стремление к истине и справедливости» [Mason & Wilson, 2017, p. 96].



Таким образом, если в случаях неявной предвзятости проявляется интеллектуальный порок, эти случаи не объясняются мотивационной теорией<sup>4</sup>.

## Являются ли эпистемические пороки неявными установками?

Для того чтобы показать, что интеллектуальные пороки могут быть неявными установками, нужно проанализировать соотношение черт характера и неявных установок. Представим такой случай:

*Скупой Барон.* Барон проявляет все признаки такой черты характера, как скупость: он не дает денег своему сыну даже на самое необходимое, совершает плохие с моральной точки зрения поступки ради денег и тому подобное. При этом Барон не признает того, что он скуп: он предпочитает называть себя бережливым, предполагая, что бережливость – это добродетель. Таким образом, он находится под действием «иллюзии бережливости». Скупость Барона проявляется автоматически, не находится под его контролем.

Можно ли сказать, что это описание скупости является нетипичным? Я склоняюсь к противоположному мнению: хотя эта черта и представлена мной несколько гротескно, однако это именно то, что мы обычно имеем в виду под скупостью. Нетрудно заметить, что в данном примере скупость имеет многие черты неявной установки: автоматизм, неподконтрольность, неосведомленность агента о наличии у него такой установки.

Для того чтобы считаться чертой характера, установка должна быть в достаточной мере укоренена в субъекте, «формировать часть того, что делает личность той личностью, которой она является» [Ibid., p. 116]. Нил Леви утверждает, что неявные установки не могут быть полностью атрибутированы агенту именно в силу их скрытости, непризнания их агентом. Однако он считает, что провал между неявными и явными установками гораздо меньше, чем иногда предполагается, и, в конце концов, атрибутирование установок агенту – это дело степени [Levy, 2017, p. 21]. Вряд ли можно найти какой-то «коэффициент укорененности», при котором установка может считаться чертой характера. По-видимому, ответ на вопрос, является ли

<sup>4</sup> Разумеется, эту теорию, как и любую другую, можно продолжить спасать любыми средствами. В конце концов, воображаемый обвинитель мог бы сказать участнику эксперимента: «При наличии достаточной мотивации ты бы не совершил ошибки, не сделал неправильного выбора». Однако вряд ли такой обвинитель мог бы заранее сформулировать критерии мотивации, достаточно сильной для того, чтобы исключить возможность проявления неявной предвзятости.



какая-то установка чертой характера агента, зависит от того, как часто и в каких условиях она срабатывает, как она связана с другими установками. При этом могут быть и спорные случаи, и такие, в которых ответ будет очевидным. Кроме того, Загзевски утверждает, что добродетели и пороки должны быть приобретенными [Zagzebski, 1996, p. 116]. Это условие, очевидно, может быть удовлетворено в случае с неявными установками. Качество приобретенности предполагает, что черты характера субъекта можно скорректировать. Тамар Гендлер, включающая неявные установки в класс праубеждений, считает, что они могут изменяться, в том числе и самим субъектом через постоянные упражнения или размышления. Всё это показывает, что эпистемические пороки могут быть неявными установками.

## **Эпистемические пороки как явные установки**

Приведенный в конце предыдущего параграфа вывод признает и Л. Загзевски: «Личность, имеющая интеллектуальные пороки, может быть не осведомлена о том, что она их проявляет в тех случаях, когда это случается. Она может даже ошибочно принять порок за добродетель, как и в случае с моральными пороками» [Ibid., p. 152]. Однако, по-видимому, Загзевски предполагает, что такая ситуация – это лишь частный случай, а более распространенными должны быть пороки как явные установки. Далее я предложу анализ эпистемического порока как явной установки.

Если субъект обладает явной установкой, то у него есть некоторое корректное убеждение относительно нее, которое он готов отстаивать до определенной степени. Иными словами, он признает наличие у себя этой установки. Можно отстаивать тезис о том, что все пороки (и моральные, и интеллектуальные) – это неявные установки в тривиальном смысле. Большинство людей не являются тем, что Кант называет дьявольским существом: «Освобождающий от морального закона, как бы злой, разум (безусловно злая воля) содержит в себе слишком много, так как этим противодействие самому закону стало бы мотивом (ведь без того или иного мотива произвол не может быть определен), и, таким образом, субъект стал бы дьявольским существом. – Но [это] не применимо к человеку» [Кант, 1994, т. VI, с. 36]. Зло само по себе не может быть мотивом для агента, считает Кант. Личность, имеющая порок как явную установку, была бы «дьявольским существом»: она должна была бы осознано принимать зло, на которое направлен порок, как цель, иметь соответствующее убеждение, быть готовой защищать свою установку. Я не исключаю возможности существования таких «дьявольских существ», но предполагаю, что, если они есть, их очень и очень мало.



Однако если мы вернемся к парадигмальным случаям неявных установок – неявному сексизму и расизму, – то увидим, что разница между явным и неявным расистом состоит не в том, что первый является дьявольским существом и открыто стремится ко злу, тогда как второй – к благу. Явному расисту свойственно открыто признавать и защищать свои убеждения, считая их правильными. Того же мы можем ожидать и от носителя явных эпистемических пороков.

Предположительно, явные интеллектуальные пороки должны встречаться реже явных моральных пороков. Дело в том, что различие явных и неявных установок во многом основано на предрасположенности субъекта признавать и отстаивать свои убеждения. Защита моральных убеждений опирается на внешние по отношению к моральному дискурсу эпистемические нормы и практики<sup>5</sup>. Поэтому защита какого-то морального порока может строиться на эпистемически порочном основании без того, чтобы сам моральный дискурс стал противоречивым. Но это не так для интеллектуальных пороков, существующих внутри познавательного дискурса. Защита эпистемического порока неизбежно должна приводить к разрушению самого дискурса. То есть субъект, имеющий интеллектуальный порок как явную установку, должен был бы иметь некогерентную систему убеждений.

По-видимому, есть одна группа интеллектуальных пороков, которые могут *до определенной степени* последовательно защищаться их носителем. Возьмем для примера такой порок, как интеллектуальное высокомерие – убеждение в собственной интеллектуальной исключительности [Baird & Calvard, 2019, p. 270–271].

Представим себе *Босса*, который считает, что он всегда прав. Босс является интеллектуально высокомерным: игнорирует даже самые убедительные доводы других людей, считает, что их интеллектуальные способности существенно ниже его. При этом Босс признает за собой такую черту характера и не считает ее пороком.

В случае Босса мы имеем дело с интеллектуальным пороком как явной установкой. Босс – не «дьявольское существо»: он открыто признает наличие у себя интеллектуального высокомерия, однако не считает, что это свойство систематически приводит к эпистемическому злу. Далее представим *Шефа*:

*Шефу* присуще интеллектуальное высокомерие, однако он сам этого не признает и явно считает интеллектуальное высокомерие пороком. Его поведение говорит о том, что он часто игнорирует даже самые убедительные доводы других людей, считает, что их интеллек-

<sup>5</sup> Я предполагаю, что эпистемические добродетели не являются моральными. Подробнее о различных точках зрения на этот счет см. [Каримов, 2019, с. 186–199].



туальные способности существенно ниже его. Когда ему указывают на явные случаи проявления высокомерия, он отказывается признавать у себя такую черту: он впадает в казуистику, выдвигает всевозможные оправдания.

Разница между случаями Босса и Шефа очевидна, и она полностью аналогична различию между явным и неявным сексистом или расистом.

Интеллектуальные пороки, которые могут наиболее успешно претендовать на статус явных установок, составляют ограниченную, хотя и достаточно широкую, группу *эгоистических пороков*. Оправдание Боссом своего высокомерия должно быть построено либо на критериях, которые предполагают его особенное отношение к себе, либо на универсальных и равно применимых ко всем. Если мы возьмем второй случай, то возможны два сценария: либо Босс оказывается прав и он действительно обладает выдающимися качествами, тогда его высокомерие в какой-то степени оправдано; либо (что более вероятно) Босс не прав, и тогда он впадает в казуистику, превращаясь в Шефа. Таким образом, чтобы Босс мог иметь эпистемический порок как явную установку, он должен быть эгоистом и применять к себе иные критерии интеллектуальных пороков, нежели к другим<sup>6</sup>. Вера в собственную уникальность может быть (и как правило) основана на убеждении в привилегированном статусе социальной группы, к которой относится субъект (как в случае с сексизмом или расизмом).

Явные эгоистические пороки, по-видимому, не так уж редки. Однако их распространение сильно ограничивается тезисом о тотальности коммуникативного знания, который гласит, что «значительная часть наших познаний получена от других, а не путем самостоятельного усилия» [Каримов, 2019, с. 315]. Очевидно, что при условии истинности этого тезиса (а я склонен его принимать), последовательная открытая защита эгоистического порока становится невозможной: субъект либо признает необходимость пересмотра своей установки, либо эта установка переходит в область неявных, а субъект впадает в определенную форму иррациональности.

Продолжая разбирать случаи возможных явных интеллектуальных пороков, мы увидим, что отстаивающий их субъект будет впадать во всё большую иррациональность. Как быть со случаями подчинения познавательных мотивов личности мотивам, например, моральным (в эту группу можно включить и корыстные мотивы)? Представим себе ситуацию, когда субъект явно пренебрегает интеллектуальной добродетелью ради достижения некоторого морального

<sup>6</sup> Этот эгоизм не нужно путать с эпистемическим эгоизмом [Каримов, 2019, с. 329–332].



блага и открыто это признает и защищает. Не будет ли тогда интеллектуальный порок его явной установкой?

Для того чтобы утверждать, что в этих ситуациях мы имеем дело именно с явными установками субъектов, нужно показать, что здесь проявляется некоторое устойчивое убеждение, которое субъект признает и готов отстаивать. Но не все случаи подобных дилемм соответствуют такому требованию. Мы иногда сталкиваемся с ситуациями выбора из двух зол, когда любой исход плох в определенном отношении. Представим себе интеллектуально добродетельного человека, попавшего в такое положение: с одной стороны, он обоснованно убежден в том, что должен поступить морально, с другой – что не должен пренебрегать интеллектуальной добродетелью. Предположим, что этот человек делает выбор в пользу морального блага. В такой ситуации действие, ведущее к эпистемическому злу, не будет проявлением порока, поскольку оно не будет выражать характер личности: ведь наш субъект по определению не склонен пренебрегать эпистемическими добродетелями. Сама возможность представить такую ситуацию говорит о том, что не во всех подобных сценариях имеет место эпистемическая порочность. Таким образом, чтобы мы могли говорить здесь об эпистемическом пороке как явной установке, нужно, чтобы субъект имел явную склонность к проявлению этого порока в ситуациях, в которых интеллектуально добродетельный человек точно проявил бы добродетель. А это уже указывает на то, что любая попытка такого субъекта защитить свою установку приведет к тому же результату, что и защита эгоистического порока: либо он будет вынужден открыто защищать порок, что приведет его к противоречию и необходимости пересмотра установки, либо установка перейдет в статус неявной.

Можно представить себе ситуации конфликта явных установок, каждую из которых субъект склонен отстаивать, субъект может проявлять предельную иррациональность и защищать интеллектуальные пороки вопреки всему. В этих случаях он так или иначе теряет когнитивный контакт с реальностью. Ситуации, в которых имеет место явный конфликт установок (например, случаи раздельного мышления,<sup>7</sup> интеллектуальной акрасии, аддикций) представляют собой отдельный интересный предмет для исследования. Однако установки, проявляемые агентами в таких сценариях, нельзя однозначно называть явными, поскольку они часто нечувствительны к доводам.

<sup>7</sup> Под раздельным мышлением, или компартиментализацией, имеется в виду «защитный механизм, при помощи которого мысли и чувства, которые кажутся конфликтующими или несовместимыми, изолируются друг от друга в разных и, как представляется, не взаимодействующих областях психики» [American Psychological Association, web]. Термин позаимствован из психоанализа.



Наконец, нужно признать, что список интеллектуальных пороков и добродетелей нельзя считать завершенным. Кроме того, разница между пороком и добродетелью может состоять в степени проявления какого-то качества личности. Может наблюдаться теоретическое несогласие относительно того, что именно можно однозначно назвать пороком, а что – добродетелью. В случае, если вы считаете какую-то черту моего характера пороком и приводите для этого сильные, но недостаточные основания, а я в ответ привожу убедительные доводы в пользу того, что это не порок, нельзя, разумеется, сделать вывод о том, что я имею порок как явную установку.

Таким образом, можно сделать вывод о том, что агент может иметь пороки как явные установки только в силу эгоизма или иррациональности (и чем выше условная степень этой иррациональности, тем труднее признать порок явной установкой). Разумеется, из всего вышесказанного не следует, что носители интеллектуальных пороков – мы сами или люди вокруг нас – не являются иррациональными эгоистами: то есть, что интеллектуальные пороки присущи им чаще в неявной, чем в явной форме. Теперь я постараюсь показать, почему мы можем принять гипотезу, что в настоящее время в современной культуре значительная часть эпистемических пороков представляет собой именно неявные установки.

Проведем аналогию с явной и неявной дискриминацией, которую я использовал как модельный пример. Сто лет назад явные дискриминационные установки по гендерному или расовому признаку были широко распространены, тогда как сегодня они встречаются гораздо реже. Это не означает что сегодня нет явных сексистов или расистов. Просто неявные проявления дискриминации с развитием культуры и искоренением явных ее проявлений становятся всё более распространенными и представляют всё более насущную проблему. Такое же положение дел, предположительно, мы можем наблюдать и в сфере познания. Да, мы можем встретить носителей явных эпистемических пороков: возможно, «дьявольских существ», часто интеллектуальных эгоистов, иногда иррациональных субъектов. Но явные установки отличаются тем, что они чувствительны к доводам. Закрепление в культуре стандартов познавательной деятельности, критического отношения к своим убеждениям должно вести к искоренению явных интеллектуальных пороков. Вероятно, по мере искоренения эти пороки будут превращаться в неявные установки. Это именно то, чего следует ожидать и что может быть в будущем подвergнуто эмпирической проверке.



## Итоги и перспективы

Гипотеза о том, что интеллектуальные пороки часто являются неявными установками, может продуктивно применяться как в теории, так и на практике. Теоретически явные и неявные пороки чаще всего могут быть легко различены. Интеллектуальные пороки, понятия как неявные установки, удовлетворяют требованию «золотой середины» Кассама. Они не сводятся ни к локальным предпочтениям агента (неявные установки систематически проявляются агентом в ситуациях определенного типа), ни к глобальным чертам характера, как, например, догматизм или предвзятость вообще. Неявная предвзятость по отношению к представителям определенной группы проявляется систематически, но при соблюдении некоторых условий. Например, ситуация выбора, в которой оказывается субъект, должна быть достаточно эластичной (как показывают эксперименты Нортонна и его коллег) и допускать возможность «казуистических» истолкований своего поведения агентом. Таким образом, неявные установки – это такие не слишком локальные и не слишком глобальные черты личности, которые идеально подходят на роль интеллектуальных (и, возможно, моральных) пороков.

Представленный в статье анализ подталкивает нас к выводу, что одна и та же неявная установка может иметь разные проявления, некоторые из которых могут быть расценены исключительно как проявления интеллектуального порока, а некоторые – как манифестации в том числе и моральных качеств агента. В литературе обсуждается вопрос о моральной ответственности агентов за неявно предвзятое поведение: некоторые авторы считают, что неявные установки обладают такими свойствами, которые исключают возможность моральной ответственности агента за них (например, [Levy, 2017]); другие предлагают свои модели моральной ответственности за неявную предвзятость (например, [Vargas, 2017]). Вопрос о моральной ответственности в данной статье не затрагивался. Однако если правы те, кто считает, что агенты могут нести моральную ответственность за неявно предвзятое поведение, то мы можем заключить, что одна и та же неявная установка является чертой и морального, и интеллектуального характера. Этот вывод может быть использован для анализа соотношения моральных и интеллектуальных пороков и добродетелей. Трактовка интеллектуальных пороков как неявных установок также открывает возможность для построения теории эпистемической ответственности с опорой на уже имеющиеся исследования моральной ответственности за неявно предвзятое поведение. Обсуждаемая гипотеза объясняет, почему интеллектуальные пороки трудно искоренить: в отличие от явных установок, которые относительно легко меняются при наличии достаточных доводов





против них, неявные установки не чувствительны к изменению убеждений агента, а их корректировка требует больших усилий и времени.

В практическом плане понимание интеллектуальных пороков как неявных установок полезно для исследования систематических пороков, которые проявляются в организациях [Baird & Calvard], могут лежать в основе лженаучных теорий и т.д.

В целом к неявным интеллектуальным порокам можно применить все наработки по неявным установкам в области философии и психологии, что позволит значительно расширить методологический аппарат исследований черт интеллектуального характера.

## Список литературы

- Кант, 1994 – Кант И. Собр. соч.: в 8 т. Т. 6. М.: Чоро, 1994. 613 с.  
Каримов, 2019 – Каримов А.Р. Эпистемология добродетелей. СПб.: Алетейя, 2019. 428 с.  
Майерс, 2010 – Майерс Д. Социальная психология. 7-е изд. СПб.: Питер, 2010. 794 с.  
Шевченко, 2020 – Шевченко С.Ю. Презирать и подсказывать: эпистемическая несправедливость и контр-экспертиза // Эпистемология и философия науки. 2020. Т. 57. № 2. С. 20–32.

## References

- Alfano, 2012 – Alfano, M. “Expanding the Situationist Challenge to Responsibility Virtue Epistemology”, *The Philosophical Quarterly*, 2012, vol. 62, pp. 223–249.  
American Psychological Association, web – American Psychological Association. “Explicit Attitude”, *APA Dictionary of Psychology*. [<https://dictionary.apa.org/explicit-attitude>, accessed on 28.02.2021].  
Baird & Calvard, 2019 – Baird, C. & Calvard, T.C. “Epistemic Vices in Organizations: Knowledge, Truth, and Unethical Conduct”, *Journal of Business Ethics*, 2019, vol. 160, no. 1, pp. 263–276.  
Brownstein, 2019 – Brownstein, M. “Implicit Bias”, in: E.N. Zalta (ed.). *The Stanford Encyclopedia of Philosophy (Fall 2019 Edition)*. 2019. [<https://plato.stanford.edu/archives/fall2019/entries/implicit-bias/>, accessed on 28.02.2021].  
Cassam, 2016 – Cassam, Q. “Vice Epistemology”, *The Monist*, 2016, vol. 99, no. 2, pp. 159–180.  
De Houwer et al., 2009 – De Houwer, J. et al. “Implicit Measures: A Normative Analysis and Review”, *Psychological Bulletin*, 2009, vol. 135, no. 3, pp. 347–368.  
Fricker, 2007 – Fricker, M. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press, 2007. 208 p.



Gendler, 2008 – Gendler, T. “Alief in Action (and Reaction)”, *Mind & Language*, 2008, vol. 23, no. 5, pp. 552–585.

Kant, 1994 – Kant, I. *Sobranie sochinenii v vos'mi tomakh* [Collected Works in 8 Vols]. Moscow: Choro, 1994, 613 p. (In Russian)

Karimov, 2019 – Karimov, A.R. *Epistemologiya dobrodeteli* [Virtue Epistemology]. Saint-Petersburg: Aleteiya, 2019, 428 pp.

Levy, 2014 – Levy, N. “Consciousness, Implicit Attitudes and Moral Responsibility”, *Noûs*, 2014, vol. 48, no. 1, pp. 21–40.

Levy, 2017 – Levy, N. “Implicit Bias and Moral Responsibility: Probing the Data”, *Philosophy and Phenomenological Research*, 2017, vol. 94, no. 1, pp. 3–26.

Levy, 2015 – Levy, N. “Neither Fish nor Fowl: Implicit Attitudes as Patchy Endorsements”, *Noûs*, 2015, vol. 49, no. 4, pp. 800–823.

Mason & Wilson, 2017 – Mason, E. & Wilson, A.T. “Vice, Blameworthiness, and Cultural Ignorance”, in: P. Robichaud & J.W. Wieland (eds). *Responsibility: The Epistemic Condition*. New York: Oxford University Press, 2017, pp. 82–100.

Myers, 2010 – Myers, D. *Sotsial'naya psikhologiya. 7-e izd.* [Social Psychology, 7th edition]. Saint-Petersburg: Piter, 2010, 794 pp.

Norton et al., 2004 – Norton, M.I., Vandello, J.A., and Darley, J.M. “Casuistry and Social Category Bias”, *Journal of Personality and Social Psychology*, 2004, vol. 87, no. 6, pp. 817–831.

Payne, 2006 – Payne, K.B. “Weapon Bias: Split-Second Decisions and Unintended Stereotyping”, *Current Directions in Psychological Science*, 2006, vol. 15, no. 6, pp. 287–291.

Schwitzgebel, 2010 – Schwitzgebel, E. “Acting Contrary to Our Professed Beliefs or the Gulf between Occurrent Judgment and Dispositional Belief”, *Pacific Philosophical Quarterly*, 2010, vol. 91, no. 4, pp. 531–553.

Schwitzgebel, 2019 – Schwitzgebel, E. “Belief”, *The Stanford Encyclopedia of Philosophy (Fall 2019 Edition)*. Edward N. Zalta (ed.), 2019. [<https://plato.stanford.edu/archives/fall2019/entries/belief/>, accessed on 28.02.2021].

Shevchenko, 2020 – Shevchenko, S.Y. “Prezirat' i podskazyvat': epistemicheskaia nespravedlivost' i kontr-ekspertiza” [Incline and Admonish: Epistemic Injustice and Counter-Expertise], *Epistemology & Philosophy of Science*, 2020, vol. 57, no. 2, pp. 20–32.

Uhlmann & Cohen, 2005 – Uhlmann, E.L. & Cohen, G.L. “Constructed Criteria: Redefining Merit to Justify Discrimination”, *Psychological Science*, 2005, vol. 16, no. 6, pp. 474–480.

Vargas, 2017 – Vargas, M. “Implicit Bias, Responsibility, and Moral Ecology”, in: D. Shoemaker (ed.). *Oxford Studies in Agency and Responsibility*. Vol. 4. New York: Oxford university Press, 2017, pp. 219–247.

Zagzebski, 1996 – Zagzebski, L.T. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge University Press, 1996, 384 pp.