

ЭПИСТЕМОЛОГИЯ И КОГНИТИВНЫЕ НАУКИ

А.А. Григорьев

Критический анализ концепции сознания доступа Неда Блока в контексте проблемы сознания в небиологических субстратах

Григорьев Александр Александрович – аспирант. Институт философии РАН. Российская Федерация, 109240, г. Москва, ул. Гончарная, д. 12, стр. 1; e-mail: rsonag@gmail.com

В статье содержится краткий обзор различных точек зрения на возможность воплощения искусственного сознания. Несмотря на то, что Нед Блок скорее отрицательно отвечает на вопрос о возможности сознания в машинах, на первый взгляд представляется, что его концепция разделения сознания доступа и феноменального сознания могла бы быть полезна для разрешения этой проблемы, поскольку предполагается, что первое может существовать без второго. Приводя пример на реальных фактах слепого зрения, Блок отмечает, что в этих случаях отсутствует не только феноменальное, но и сознание доступа, поскольку человека необходимо каждый раз просить догадываться, какой предмет перед ним находится. Чтобы рассмотреть работу сознания доступа без феноменального сознания, следовало бы представить «суперслепое зрение», когда испытуемый дискриминировал бы предметы в пространстве самостоятельно. Однако если такое «сознание» может существовать без субъективного, качественного опыта, то оно либо не отличается от механической (бессознательной) переработки информации, либо, напротив, обладает всем богатством феноменального сознания, когда субъект сам видит все предметы и не требуется их угадывать. Таким образом, сознание доступа не является чем-то самостоятельным и достаточным для определения сознания. При этом пример со слепым зрением может служить наглядной иллюстрацией и аналогией того, каким образом осуществляется регулирование обработки информации, к примеру, в больших языковых моделях, в которых возможно влиять на результаты получаемых ответов через температуру или изменение заданных параметров (изменение степени влияния). Однако сама по себе доступность информации в случаях слепого зрения не является автономным сознанием, это лишь следствие тонкой настройки модели за счет изменения ее параметров. Эти выводы подчеркивают важность феноменального аспекта сознания для рассмотрения возможности его воплощения в небиологических субстратах.

Ключевые слова: феноменальное сознание, сознание доступа, философия сознания, искусственное сознание, слепое зрение, небиологические субстраты, общий искусственный интеллект, температура, большая языковая модель

Введение

Различие двух типов сознания – феноменального сознания и сознания доступа – было предложено философом Недом Блоком. Н. Блок иллюстрирует понятие сознания доступа на примере феномена слепого зрения, когда у человека повреждена зрительная зона головного мозга V1 и он определяет себя полностью незрячим, тем не менее дискриминирует предметы в пространстве, и с увеличением частоты тренировок растет вероятность их правильного распознавания. Из этого Н. Блок заключает, что у человека хотя и отсутствует компонент феноменального сознания, ответственный за зрительное восприятие, тем не менее сознание доступа продолжает получать информацию о предметах, при этом сам человек как будто догадывается об этом [Block, 1995].

В более поздней работе Н. Блок приводит обратный пример, когда у младенца в возрасте от 6 до 11 месяцев при распознавании цветов задействовано феноменальное сознание, он проявляет различные реакции при изменении цвета, но, как полагает Н. Блок, у него отсутствует сознание доступа, поскольку еще не сформировалась «концепция цвета». При этом сознание доступа представляет собой трансляцию для свободного использования в рассуждениях и для прямого рационального контроля действий [Block, 2023].

Возможно ли использовать такое разделение сознания для работы над вопросом воплощения искусственного сознания? Ответ скорее отрицательный, и вот почему.

Приводя пример на реальных фактах слепого зрения, Н. Блок отмечает, что в этих случаях отсутствует не только феноменальное, но и сознание доступа, поскольку человека необходимо каждый раз просить догадаться, что перед ним находится, сам он без команды не будет этого делать и «не потянется за стаканом воды на столе», даже если испытывает жажду. Чтобы рассмотреть работу сознания доступа без феноменального сознания, следовало бы предположить возможность суперслепого зрения, когда «незрячий» человек дискриминировал бы предметы в пространстве без того, чтобы его просили делать это (догадываться).

Однако, определяя таким образом сознание доступа и отделяя его от феноменального сознания, мы неизбежно впадаем в противоречие относительно того, что считать сознанием и что в таком случае выступает предметом исследования.

Различные подходы к возможности воплощения искусственного сознания

На функциональном предназначении сознания и потенциальной возможности его появления у машин настаивает французский нейробиолог, развивающий теорию глобального рабочего пространства (GWT) Станислас Деан.

Однако Нед Блок, Кристофер Кох и другие известные ученые не соглашаются с этим, утверждая, что машина не в силах обладать сознательным опытом, не испытывая качественных переживаний. По их мнению, несмотря на то, что искусственный интеллект пока не может продемонстрировать трансляцию информации в рабочее пространство или формально соответствовать правилам байесовского вывода, уже сейчас многие системы включают глобальную проекцию выигрышных представлений и обладают метакогнитивными способностями. Если такие типы процессов свидетельствовали бы о сознании, необходимо заключить, что некоторые машины уже им обладают [Carter, Hohwy et al., 2018, p. 400].

Деан отвечает: «Мы далеки от консенсуса в понимании сознания, именно поэтому необходимо обращаться к эмпирическим данным и может оказаться так, что феноменальный опыт представляет собой всего лишь информационные репрезентации с теми же атрибутами, которые обрабатываются машинами: глобальная доступность для отчета и использование в определенных процессах самоконтроля, способных к оценке и критике. Так, при наблюдении нарушений, связанных с феноменальным сознанием, к примеру, психоза или слепого зрения, обе вышеупомянутые функции также страдают. Из этого следует рабочая гипотеза, что приватный опыт не что иное, как комбинация специальных форм обработки информации, включая мониторинг реальности как ключевой компонент». В итоге Деан заключает: «когда машины будут обладать достаточными функциональными характеристиками, сходными с функциями сознательной обработки информации человеком, нам следует быть готовыми, чтобы признать их осознанными». Отрицание этого факта, по мысли Деана, «схоже с солипсистской позицией, ровным счетом применимой к людям: по такому стандарту мы также не можем доказать, что другие люди обладают сознанием» (перевод мой. – А.Г.), но такой подход является явно контрпродуктивным [Ibid., 2018, p. 401].

На невозможности определения четкой границы между сознательным и бессознательным настаивал американский философ Дэниел Деннет. Он подчеркивал нейрофизиологическую природу сознания, его нестабильность и зависимость от разных факторов. Большая часть наших мыслей и решений принимается на уровне подсознания, а потом приписывается сознательному «я», в подтверждение чему приводились результаты различных научных экспериментов [Dennett, 1991].

Д. Деннет утверждал, что системы искусственного интеллекта, если они разработаны и реализованы с использованием правильных вычислительных структур, могут обладать сознательным опытом. При этом наше современное понимание сознания находится под неоправданным влиянием идей Рене Декарта. Сознание – это иллюзия, и феноменального опыта попросту не существует [Ibid.].

Более центристский взгляд на эту проблему у Дэвида Чалмерса – хотя он и формулирует трудную проблему сознания, заявляя о несводимости ментального опыта только к физическим процессам и сомневаясь в выборе между дуалистическими и панпсихистскими гипотезами, тем не менее приходит к выводу: «В самом деле, можно было бы попробовать показать, что если каждый

элемент мозга, такой как нейрон, имеет лишь конечное множество релевантных состояний и если существует лишь конечное множество релевантных элементов, то релевантная каузальная структура мозга должна выражаться вычислительным описанием» [Chalmers, 1996], а раз так, то возникновение сознания при перенесении соответствующего вычислительного описания вполне допустимо в искусственном интеллекте.

В современной отечественной философии примечателен подход И.Ф. Михайлова: «Очевидно, что т.н. “чувственный опыт” – это биологическая форма доступа к данным. В случае с машиной, наделенной сознанием, форма доступа к данным может быть самой разной, в зависимости от потребности создателя: доступ к уже сформированной базе данных, имитация биологических органов чувств, ультразвуковое или инфракрасное сканирование предметной области или что-либо еще. Главное состоит в необходимости воспроизведения блок-схемы: потребность → множество возможных значений → значащий чувственный образ (← язык) ⇒ онтология (Umwelt)» [Михайлов, 2012, с. 55].

Таким образом, в современной философии и науке, несмотря на разность методологических подходов, постепенно укореняется мысль о принципиальной возможности воплощения искусственного сознания. Однако на этом пути нам следовало бы определиться с самим понятием сознания или, по крайней мере, с предметом для искусственного воплощения, в чем, несомненно, возникает масса трудностей. Одним из подступов к этому вопросу как раз и представлялось предложенное Н. Блоком разделение: сознание доступа и феноменальное сознание. Может ли так получиться, что воплощения искусственного сознания доступа будет достаточно для утверждения о существовании искусственного сознания? Более того, может быть, искусственное сознание доступа уже существует? Далее попробуем наметить пути для того, чтобы разобраться в этих вопросах.

Доступность информации как неотделимое свойство феноменального сознания

Согласно Н. Блоку, состояние «доступно – сознательно», если его содержание готово к использованию в качестве посылки при размышлении для рационального контроля действия и речи. Так что сознание доступа приблизительно соответствует дефиниции осведомленности (осознанности) Д. Чалмерса, хотя в ней меньше акцентируется рациональность [Chalmers, 1996]. Осведомленность (осознанность) Д. Чалмерс определяет как «прямую доступность для глобального контроля. То есть субъект осведомлен о какой-то информации, если эта информация напрямую доступна для осуществления широкого спектра поведенческих процессов. Это позволяет говорить о возможности опытных переживаний у животных, отличных от человека, и хорошо сочетается с критерием отчета. В случаях, где об информации можно отчитаться, она, как правило, доступна для глобального контроля (в частности, для намеренной координации широкого спектра поведенческих актов)» [Ibid.].

Из этого можно заключить, что доступность информации для рационального (глобального) контроля выводит ее на осознаваемый уровень. Однако

данная информация обычно переживается именно в феноменальном смысле, каким-то образом представляясь субъекту для возможности ее дальнейшей сознательной обработки. Сознание доступа в таком понимании в общем случае выступает как бы интегратором для феноменальных переживаний и работает вместе с феноменальным сознанием. Но если мы попытаемся отделить одно от другого, например, представив суперслепое зрение, то нам либо ничего не останется от осознанного опыта или же, наоборот, мы будем обладать, по сути, богатым феноменальным опытом.

Попытка классифицировать таким образом сознание, определив процесс интеграции и «вывода на дисплей» сенсорного восприятия в качестве типа сознания, приводит к размыванию самого понятия и рассуждению о том, что само по себе не осознается и сознанием в буквальном смысле не является. Ведь даже рассматривая примеры наличия одного сознания без другого, Н. Блок различает для каждого типа сознания свой вид зомби, которым было бы представлено существо, обладающие только одним типом сознания. Само по себе это разграничение уже говорит о том, что одно без другого не работает как сознание, а следовательно, в практическом смысле такое разделение несостоятельно.

Не исключая слепого зрения и других примеров получения доступа к информации, возникновение этой информации в сознании в виде некого образа уже приобретает феноменальные свойства. Попытка представить сознание доступа как данное в репрезентациях, а феноменальное сознание только как представление квалиа, не решает эту проблему, поскольку репрезентации лишь отражают квалиа.

При этом существует *буквальное противоречие* в понимании сознания доступа, объясняемого через пример со слепым зрением, когда нечто недоступное осознанному контролю влияет на поведение, и совершенно противоположное ему свойство сознания, когда интегрируется многообразие феноменального опыта, становясь доступным для контроля и осознанных действий.

Для вывода о том, что поведение человека во многом детерминировано неконтролируемыми факторами, совсем необязательно приводить пример со слепым зрением. Львиная доля различных физиологических реакций так или иначе обусловлена факторами, лежащими за пределами осознания, но которые могут являться как нечто осознаваемое и переживаемое в опыте. Возьмем любой физиологический позыв: будь то чувство голода или жажды, влечение или усталость, все это так или иначе вызывается имплицитными химическими процессами, которые репрезентируются в виде этих чувств, являясь своего рода суррогатами этой скрытой от сознания информации. О том, что мы воспринимаем мир всегда как нечто производное, говорил еще Д. Юм [Hume, 2007].

Определяя так сознание доступа (через пример слепого зрения), можно было бы прийти к выводу, что все имплицитные процессы в организме и во внешней среде, которые влияют на организм, но при этом буквально не выражены в феноменальном опыте, начиная от взаимодействия клеток в организме и заканчивая социальными процессами, порождая информацию, которая в той или иной мере влияет на поведение субъекта, являются собой сознание

доступа. Только проблема в том, что все эти процессы скрыты от субъекта и им буквально не осознаются, а следовательно, слово сознание здесь попросту является лишним.

Если даже признать наличие сознания доступа, то его рассмотрение в отрыве от феноменального сознания практически ничего не привнесет в понимание возможности воплощения искусственного сознания в небиологических субстратах, поскольку идеальным примером наличия сознания доступа без феноменального сознания, который приводит сам же Н. Блок, является зомби, который может обрабатывать информацию и ориентироваться в окружающей среде, не имея ни малейшего представления об этом. И, по всей видимости, такие машины уже существуют.

Если исходить из того, что сознание доступа заключается в получении какой-то репрезентации для ее обработки и, как следствие, использования для действия (поведения), то следует признать, что любая программа, получающая ту или иную входящую информацию, обрабатывающая ее функционально и выдающая определенный результат, обладает таким сознанием. По крайней мере, большие языковые модели имеют возможность давать отчет о той информации, которой они располагают, используя ее в своих рассуждениях. К примеру, ChatGPT фактически признает наличие у него всех свойств, которые приписываются сознанию доступа.

Также неверно было бы и отказывать феноменальному сознанию в функциональности, рассматривая его исключительно как поток информации, или как поток воды в гидравлическом компьютере, как представляет его Н. Блок, поскольку обработка и представление информации в воспринимаемом субъектом виде уже есть переработка входящих данных из среды. Таким образом, сознание доступа ни из чего не выводится и неотделимо от феноменального сознания.

Утверждение о том, что лунатик, объезжая препятствия, обладает феноменальным сознанием, к которому нет доступа, так же не фальсифицируемо, как и утверждение, что он обладает «полным» сознанием, но забывает все после пробуждения или что он делает все на автоматизмах, не обладая сознанием в его феноменальном аспекте. Например, опытный водитель может не помнить дорогу и то, как приехал домой, разговаривая по телефону. Если попросить такого человека вспомнить путь, то он навряд ли это сделает, потому что дорога не была предметом его рассмотрения. А значит, у него не было сознания доступа? Или сознание доступа «не подключилось» к феноменальному? Или, может быть, такое разделение было бы неверным, поскольку сознание человека не было направлено на определенный объект, а было занято другими делами? Тогда как, попав в небольшую аварию, он переживал бы этот момент во всех красках, но, если бы он попал в крупное происшествие, из памяти стерся бы момент аварии. В этой связи однозначно нельзя утверждать, к кому ближе водитель, который не помнит дорогу: к тому, кто попал в аварию и потерял память, но имел феноменальный опыт, или к Яндекс-роботу, который объезжает препятствия, но, как нам пока представляется, все-таки не имеет феноменального сознания. Однако ясно одно, что, даже решив эту задачу, мы не приблизились бы к решению вопроса о возможности воплощения искусственного

сознания, поскольку ни в том, ни в другом случае такой опыт не будет в полном смысле осознаваемым.

Пример слепого зрения можно проиллюстрировать совершенно в других терминах: богатство содержания (richness of content) и степень влияния (degree of influence) [Dennett, 1995]. Так, сам по себе стакан воды, стоящий на столе перед «незрячим», несмотря на его жажду, из-за бедности содержания такого опыта не будет способен вызвать стимул, чтобы потянуться за ним и утолить жажду. Однако если мы попросим угадать предмет, стоящий перед ним, т.е. намеренно направим его внимание, тем самым увеличив степень влияния на его опыт как бы менее значимых, а следовательно, неосознаваемых явлений, то в таком случае он при должной степени натренированности с хорошей вероятностью распознает, что перед ним стакан воды.

Такое положение дел напоминает, каким образом может регулироваться вероятность того или иного ответа (вывода) в больших языковых моделях при изменении параметров. Присвоение весов последовательности символов в ответе исходит из того контента, который рассматривается при заданных начальных параметрах, т.е. на основе символов, заданных в вопросе. Просьба угадать предмет, которого нет в поле зрения смотрящего и содержание которого крайне бедно для него и без этой просьбы не определялось бы, представляется полным изменением входящих параметров и максимальным увеличением степени влияния предмета, влияние которого крайне мало при обычных условиях из-за бедности его содержания. То есть одним из определяющих факторов степени влияния является богатство содержания само по себе, но это не единственный фактор, поскольку изменение степени влияния того или иного содержания может происходить из-за воздействия на модель как внешних (эксплицитных), так и внутренних (имплицитных) факторов. Как было показано выше, одним из таких внешних факторов может быть изменение запроса (просьба угадать объект), а следовательно, пересборка весов из других параметров контента и степени их влияния.

К внутренним факторам модели относится так называемая температура, которая влияет на вероятность выдачи большой языковой моделью менее ожидаемых (вероятных) ответов. Повышение температуры приводит к тому, что менее ожидаемым вариантам ответов, которые при низкой температуре не принимались бы во внимание, присваиваются веса и повышается вероятность их выдачи. Таким образом, температура выступает своего рода регулятором степени влияния на ответы языковой модели, которая при низкой температуре отдает предпочтение наиболее детерминированному богатством содержания (количеством встречающихся в опыте повторений) варианту, а при повышении температуры увеличивается вероятность выдачи необычных ответов. Такая ситуация в большей степени напоминает суперслепое зрение, когда испытуемого уже не просят угадать предмет перед ним, а он дискриминирует предметы в пространстве самостоятельно. Однако и в этом случае мы приходим к выводу о практической нереализуемости суперслепого зрения, как оно представляется Н. Блоком, поскольку определение предметов вне зоны видения испытуемого становится одним из множества возможных вариантов ответов (действий) при заданных параметрах и вероятность выдачи в ответе видимого нами предмета все равно остается крайне невысокой.

Таким образом, пример со слепым зрением может служить наглядной иллюстрацией того, каким образом осуществляется регулирование обработки информации и на каких основаниях она становится доступной для глобального контроля. Однако сама по себе доступность информации в таких случаях в том или ином виде не является автономным сознанием, это лишь следствие тонкой настройки модели за счет изменения ее внешних и внутренних параметров. При этом именно в феноменальном сознании следует искать функциональные свойства обработки и представления осознаваемой информации. Следовательно, при изучении возможности воплощения искусственного сознания выделение сознания доступа как отдельного типа сознания представляется непродуктивным.

Список литературы

Васильев, 2009 – *Васильев В.В.* Трудная проблема сознания. М.: Прогресс-Традиция, 2009. 269 с.

Гусев, 2021 – *Гусев А.А.* Понятие «квалиа» в репрезентационалистских теориях сознания: автореф. дис. ... канд. филос. наук: 09.00.01. М., 2021. 23 с.

Кант, 1994 – *Кант И.* Критика чистого разума / Пер. с нем. Н. Лосского. М.: Мысль, 1994. 591 с.

Кузнецов, 2015 – *Кузнецов А.В.* Проблема феноменального сознания и сознания доступа. Пример взаимодействия когнитивной науки и философии // Актуальные вопросы нейрофилософии: Материалы международного междисциплинарного семинара «Нейрофилософия», 2014–2015 гг., г. Москва / Ред. А.Ю. Алексеев, Д.И. Дубровский, В.Г. Кузнецов. М.: ИИнтелл, 2015. С. 250–257.

Михайлов, 2012 – *Михайлов И.Ф.* «Искусственный интеллект» как аргумент в споре о сознании // Эпистемология и философия науки. 2012. Т. 22. № 2. С. 107–122.

Михайлов, 2021 – *Михайлов И.Ф.* Когнитивные основания социальности: дис. ... докт. филос. наук: 09.00.01. М., 2021. 340 с.

Пшенокова, 2022 – *Пшенокова И.А.* Основные методы и подходы к моделированию искусственного сознания // Известия Кабардино-Балкарского научного центра РАН. 2022. № 2 (106). С. 72–80.

Шопенгауэр, 2023 – *Шопенгауэр А.* Мир как воля и представление / Пер. с нем. Ю.И. Айхенвальда. М.: АСТ, 2023. 672 с.

Block, 1995 – *Block N.* On a confusion about a function of consciousness // Behavioral and brain sciences. 1995. Vol. 18. No. 2. P. 227–247.

Block, 2005 – *Block N.* Two neural correlates of consciousness // TRENDS in Cognitive Sciences. 2005. Vol. 9. No. 2. P. 269–270.

Block, 2023 – *Block N.* The Border Between Seeing and Thinking. Oxford: Oxford University Press, 2023. 542 p.

Carpenter, Grossberg, 1998 – *Carpenter G.A., Grossberg S.* Adaptive resonance theory // The Handbook of Brain Theory and Neural Networks / Ed. by M.A. Arbib. Cambridge, Massachusetts: MIT Press, 1998. P. 79–82.

Carter, Hohwy et al., 2018 – *Carter O., Hohwy J., van Boxtel J., Lamme V., Block N., Koch C., Tsuchiya N.* Conscious machines: Defining questions // Science. 2018. Vol. 359. No. 6374. P. 400–402.

Chalmers, 1996 – *Chalmers D.J.* The Conscious Mind: In Search of a Fundamental Theory. New York: Oxford University Press, 1996. 414 p.

- De Quinsey, 2006 – *De Quinsey C.* Switched-on Consciousness. Clarifying What It Means // Journal of Consciousness Studies. 2006. Vol. 13. No. 4. P. 7–12.
- Dennett, 1991 – *Dennett D.C.* Consciousness Explained. Boston: Little, Brown and Co., 1991. 511 p.
- Dennett, 1995 – *Dennett D.C.* The path not taken // Behavioral and brain sciences. 1995. Vol. 18. No. 2. P. 252–253.
- Dennett, web – *Dennett D.C.* Dangerous Ideas // Tufts University. 01.09.2020. URL: <https://youtu.be/v9dZeJ3owGg> (дата обращения: 15.08.2023).
- Gazzaniga, 2018 – *Gazzaniga M.S.* The Consciousness Instinct: Unraveling the Mystery of How the Brain Makes the Mind. New York: Farrar, Straus and Giroux, 2018. 328 p.
- Hume, 2007 – *Hume D.* A Treatise of Human Nature. Oxford: Oxford University Press, 2007. 736 p.
- Jackson, 1982 – *Jackson F.* Epiphenomenal Qualia // The Philosophical Quarterly. 1982. Vol. 32. No. 127. P. 127–136.
- Levine, 1983 – *Levine J.* Materialism and Qualia: The Explanatory Gap // Pacific Philosophical Quarterly. 1983. Vol. 64 (4). P. 354–361.
- Mikhailov, 2022 – *Mikhailov I.F.* Philosophy and Science on the Way to Knowing and Making Consciousness // RUDN Journal of Philosophy. 2022. Vol. 26. No. 4. P. 740–754.
- Nagel, 1974 – *Nagel T.* What is it like to be a bat? // Philosophical Review. 1974. Vol. 83. P. 435–450.
- Penrose, 1994 – *Penrose R.* Shadows of the mind: A search for the missing science of consciousness. Oxford: Oxford University Press, 1994. 480 p.
- Rosenthal, 1986 – *Rosenthal D.M.* Two concepts of consciousness // Philosophical Studies. 1986. Vol. 49. No. 3. P. 329–359.
- Searle, 1980 – *Searle J.R.* Minds, brains, and programs // Behavioral and Brain Sciences. 1980. Issue 3. P. 417–424.
- Searle, 2015 – *Searle J.R.* Seeing things as they are: a theory of perception. Oxford: Oxford University Press, 2015. 256 p.

A critical analysis of Ned Block’s access consciousness concept in the context of the consciousness problem in non-biological substrates

Aleksandr A. Grigoriev

RAS Institute of Philosophy. 12/1 Goncharnaya Str., 109240, Moscow, Russian Federation; e-mail: rsonag@gmail.com

The article provides a brief overview of various points of view on the possibility of embodying artificial consciousness. Although Ned Block’s answer to the question of the possibility of consciousness in machines is rather negative, at first glance, it seems that his concept of the division of consciousness into access consciousness and phenomenal consciousness could help resolve this problem since it is assumed the former can exist without the second. Giving an example based on real facts of blindsight, Blok notes that in these cases, there is not only a phenomenal but also an access consciousness since a person must be asked every time to guess what object is in front of him. To consider the operation of access consciousness without phenomenal consciousness, one would imagine “super-blindsight”, where the subject would discriminate objects in space independently. However, if such “consciousness” can exist without subjective, qualitative experience, then it either does not differ from the mechanical (unconscious) processing of information or, on the contrary, we have all the wealth phenomenal consciousness when the subject himself sees all objects and does not

need to guess them. Thus, access consciousness is not independent and sufficient to define consciousness. This example of blindsight provides an obvious illustration and analogy of how information processing is regulated, for example, in large language models, in which it is possible to influence the results of the resulting reactions through temperature or changes in set parameters (changes in the degree of protection). However, the availability of information in cases of blind vision is not only an autonomous consciousness; it requires fine-tuning the model to change its parameters. These findings emphasize the phenomenal aspect of consciousness to consider the possibility of its embodiment in non-biological substrates.

Keywords: phenomenal consciousness, access consciousness, philosophy of consciousness, artificial consciousness, blindsight, non-biological substrates, artificial general intelligence (AGI), temperature, large language model

References

- Block, N. *The Border Between Seeing and Thinking*. Oxford: Oxford University Press, 2023. 542 pp.
- Block, N. "On a confusion about a function of consciousness", *Behavioral and Brain Sciences*, 1995, vol. 18, no. 2, pp. 227–247.
- Block, N. "Two neural correlates of consciousness", *TRENDS in Cognitive Sciences*, 2005, vol. 9, no. 2, pp. 269–270.
- Carter, O., Hohwy, J., van Boxtel, J., Lamme, V., Block, N., Koch, C., Tsuchiya, N. "Conscious machines: Defining questions", *Science*, 2018, vol. 359, no. 6374, pp. 400–402.
- Chalmers, D.J. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press, 1996. 414 pp.
- Dennett, D.C. *Consciousness Explained*. Boston: Little, Brown and Co, 1991. 511 pp.
- Dennett, D.C. "Dangerous Ideas", *Tufts University*, 01.09.2020. URL: <https://youtu.be/v9dZeJ3owGg> (accessed on: 15.08.2023).
- Dennett, D.C. "The path not taken", *Behavioral and Brain Sciences*, 1995, vol. 18, no. 2, pp. 252–253.
- De Quinsey, C. "Switched-on Consciousness. Clarifying What It Means", *Journal of Consciousness Studies*, 2006, vol. 13, no. 4, pp. 7–12.
- Carpenter, G.A., Grossberg, S. "Adaptive resonance theory", in: *The Handbook of Brain Theory and Neural Networks*, ed. by M.A. Arbib. Cambridge, Massachusetts: MIT Press, 1998, pp. 79–82.
- Gazzaniga, M.S. *The Consciousness Instinct: Unraveling the Mystery of How the Brain Makes the Mind*. New York: Farrar, Straus and Giroux, 2018. 328 pp.
- Gusev, A.A. *Ponyatie "kvalia" v reprezentatsionalistskikh teoriyakh soznaniya: avtoref. dis. ...kand. filos. nauk: 09.00.01*. [The concept of "qualia" in representational theories of consciousness: abstract of the diss. ... CSc in Philosophy: 09.00.01]. Moscow, 2021. (In Russian)
- Hume, D. *A Treatise of Human Nature*. Oxford: Oxford University Press, 2007. 736 pp.
- Jackson, F. "Epiphenomenal Qualia", *The Philosophical Quarterly*, 1982, vol. 32, no. 127, pp. 127–136.
- Kant, I. *Kritika chistogo razuma* [Critique of Pure Reason], trans. from German by N. Losky. Moscow: Mysl' Publ., 1994. 591 pp. (In Russian)
- Kuznetsov, A. "Problema fenomenal'nogo soznaniya i soznaniya dostupa. Primer vzaimodeistviya kognitivnoi nauki i filosofii" [The problem of phenomenal consciousness and access consciousness. An example of the interaction of cognitive science and philosophy], in: *Aktual'nye voprosy neurofilosofii: Materialy mezhdunarodnogo mezhdistsiplinarnogo seminara "Neurofilosofiya"*,

2014–2015 gg., g. Moskva [Topical issues of neurophilosophy: Materials of the international interdisciplinary seminar “Neurophilosophy”, 2014–2015, Moscow], A.Yu. Alekseev, D.I. Dubrovsky, V.G. Kuznetsov (eds.). Moscow: IIntell Publ., 2015, pp. 250–257. (In Russian)

Levine, J. “Materialism and Qualia: The Explanatory Gap”, *Pacific Philosophical Quarterly*, 1983, vol. 64 (4), pp. 354–361.

Mikhailov, I.F. “‘Iskusstvennyy intellekt’ kak argument v spore o soznanii” [“Artificial intelligence” as an argument in the consciousness debate], *Epistemologiya i filosofiya nauki* [Epistemology and Philosophy of Science], 2012, vol. 22, no. 2, pp. 107–122. (In Russian)

Mikhailov, I.F. *Kognitivnye osnovaniya sotsial’nosti: Dis. ... dokt. filos. nauk: 09.00.01* [Cognitive foundations of sociality: Diss. ... DSc in Philosophy: 09.00.01]. Moscow, 2021. 340 pp. (In Russian)

Mikhailov, I.F. “Philosophy and Science on the Way to Knowing and Making Consciousness”, *RUDN Journal of Philosophy*, 2022, vol. 26, no. 4, pp. 740–754.

Nagel, T. “What is it like to be a bat?”, *Philosophical Review*, 1974, vol. 83, pp. 435–450.

Penrose, R. *Shadows of the mind: A search for the missing science of consciousness*. Oxford: Oxford University Press, 1994. 480 pp.

Pshenokova, I.A. “Osnovnye metody i podkhody k modelirovaniyu iskusstvennogo soznaniya” [The main methods and approaches to modeling artificial consciousness], *Izvestiya Kabardino-Balkarskogo nauchnogo tsentra RAN* [News of the Kabardino-Balkarian Scientific Center of the RAS], 2022, no. 2 (106), pp. 72–80. (In Russian)

Rosenthal, D.M. “Two concepts of consciousness”, *Philosophical Studies*, 1986, vol. 49, no. 3, pp. 329–359.

Searle, J.R. “Minds, brains, and programs”, *Behavioral and Brain Sciences*, 1980, issue 3, pp. 417–424.

Searle, J.R. *Seeing things as they are: a theory of perception*. Oxford: Oxford University Press, 2015. 256 pp.

Vasiliev, V.V. *Trudnaya problema soznaniya* [The Hard Problem of Consciousness]. Moscow: Progress-Tradition Publ., 2009. 269 pp. (In Russian)

Schopenhauer, A. *Mir kak volya i predstavlenie* [The World as Will and Representation], trans. from German by Yu.I. Eichenwald. Moscow: AST Publ., 2023. 672 pp. (In Russian)