

И.Ф. Михайлов

Квалиа, репрезентации и магия философских комнат*

Михайлов Игорь Феликсович – кандидат философских наук, старший научный сотрудник. Институт философии РАН. Российская Федерация, 109240, г. Москва, ул. Гончарная, д. 12, стр. 1; e-mail: ifmikhailov@iph.ras.ru

Проблема качественно своеобразных и невыразимых аспектов феноменального опыта – квалиа – составляет, согласно распространенному мнению, основу «трудной проблемы сознания». Однако несложный анализ роли квалиа в восприятии и идентификации произведений искусства показывает, что феноменальные образы суть лишь один из строительных материалов комплексных когнитивных и их функциональная роль различается от случая к случаю. В статье приводятся аргументы против антифизикалистской интерпретации квалиа, восходящей к знаменитой «комнате Мэри» Фрэнка Джексона. Показывается, что при детальном рассмотрении этого мысленного эксперимента ни эпистемологическое видение описываемой ситуации (квалиа как репрезентации чего-либо нефизического в мире), ни онтологическое ее видение (квалиа как собственно нефизическая сущность) нельзя признать достаточно обоснованными. При помощи концептуального анализа и мысленных экспериментов в статье показывается, что сильный репрезентационализм в философском понимании квалиа нерелевантен, а слабый репрезентационализм лишен необходимой объяснительной силы. В любом случае в феноменальных субъективных образах усматривается нерепрезентативный «остаток», который играет важную каузальную или функциональную роль в человеческих восприятиях. В то же время умеренный репрезентационализм уместен и перспективен в рамках когнитивной науки, если рассматривать репрезентации как необходимый элемент вычислительных процессов где бы то ни было, в том числе в человеческом когнитивном аппарате. Однако очевидно, что в таком виде это понятие уже не претендует на роль универсального ключа для философских загадок, и его «импорт» в философию предполагает отказ от интенциональной интерпретации репрезентативного характера квалиа.

Ключевые слова: квалиа, сознание, репрезентация, репрезентационализм, интенциональность, когнитивная наука, вычисления

* Исследование выполнено при поддержке гранта Российского научного фонда, проект № 16-18-10229 «Проблема субъективности в современном междисциплинарном контексте взаимодействия философии и когнитивных наук».

1. Когда проблема квалиа действительно важна

В музее, известном как Лувр, что находится в городе, известном как Париж, выставлена картина Леонардо да Винчи, известная как «Портрет госпожи Лизы дель Джокондо». Если вам удастся протолкнуться сквозь толпы туристов, вы увидите фигуру женщины с темно-русыми волосами, в зеленом складчатом платье с желтыми рукавами, сидящую на фоне зеленовато-голубовато-розоватого пейзажа. Все, что составляет качественные характеристики этого, по мнению многих, величайшего произведения изобразительного искусства, навсегда останется у вас в памяти и будет время от времени вставать перед вашим мысленным взором с разной степенью детальности.

Но вот Лувр закрылся на ночь, вы пошли гулять по вечерним Елисейским полям, а холст остался в полутемном и пустом помещении, вдали от каких бы то ни было глаз. Где в этот момент существует картина с бессмертным образом третьей жены удачливого флорентийского торговца тканями – в музее, в вашей (и не только) зрительной памяти или нигде?

Позитивно мыслящий человек, привыкший полагаться на здравый смысл, удивится этому вопросу и увидит в нем еще один повод посмеяться над философами: наш человеческий жизненный мир (Umwelt) устроен таким образом, что картины остаются на своих местах в музеях, пока их не украдут или не отдадут на реставрацию. А вот, например, физик или химик уже может задуматься. Ведь, глядя на «Мону Лизу», мы видим не форму органических молекул красок и волокон холста, которые только и существуют объективно согласно нашей научной картине мира, фактически наследующей локковское различие первичных и вторичных свойств. Мы видим «субъективный» образ, который как раз в музее не остается, а в несколько упрощенном виде уносится нами с собой. Леонардо создал некую физико-химическую матрицу, которая при столкновении с нашим когнитивным аппаратом порождает зрительную иллюзию, целиком состоящую из того, что философы называют квалиа. Что же именно образует великое произведение искусства – «объективная» матрица или «субъективная» иллюзия?

Картина как когнитивно-социальный комплекс имеет сложную многоуровневую онтологическую структуру. Какие-то элементы и уровни этой структуры можно отнести к достаточным условиям, какие-то – к необходимым. Очевидно, что наличие качественных восприятий, например цвета, относится к необходимым условиям существования картины как картины: без них произведение живописи в отличие, скажем, от литературного произведения немислимо. Следовательно, без когнитивных ресурсов, поставляемых мозгом, например цветовосприятия, картины нет. Но ее также нет и без сложного комплекса коммуникативно интерпретируемых ресурсов, которые можно отнести к социально-когнитивным: миф, общепринятая семантика символов, эмоциональные паттерны, такие как выражение лица, и многое другое, что порождается не нейронной, а социальной сетью.

Если это было бы не так, то поздно ночью, когда Лувр закрыт, «Мона Лиза» переставала бы существовать. Но она продолжает существовать как социально-интенциональный объект в коллективной памяти, в коммуникационных потоках, ссылках и т. п. В случае (не дай бог!) утраты ее можно восстановить

в физическом воплощении с той или иной степенью достоверности, подобно тому, как были восстановлены разрушенные войной Гданьск и Варшава. Я бы назвал это случаем двойной интенциональности: картина как изображение указывает на некий интенциональный объект, чье существование предполагается за пределами холста. Но и сама она существует как интенциональный объект в коллективном сознании, в какой-то степени независимый от индивидуальной материи. Например, при реставрации мы не всегда отдаем себе отчет, да и не всегда интересуемся, какая часть материальных условий (те же холст и краски) была обновлена. Мы воспринимаем этот культурно-исторический объект как *тот же самый*.

А теперь представим себе ситуацию, когда в результате серии последовательных реставраций, осуществленных по всем правилам, с сохранением состава материалов, методов и рецептов соответствующей эпохи, физическая основа заменяется полностью. Тогда поздно ночью в Лувре висит уже другой холст с другими нанесенными на него красками. А картина – та же или другая?

Если она та же – а таковым, как мне кажется, будет ответ большинства, – тогда возникает вопрос о природе идентичности культурного объекта. Наиболее простой и очевидный ответ на этот вопрос таков: культурный объект, предполагающий его непосредственное чувственное качественное восприятие (например, картина), идентифицируется как интенциональный объект. То есть он остается *тем же самым*, пока некое сознание – индивидуальное или коллективное – не видит отличий, достаточных для того, чтобы идентифицировать его как *другой*, и, следовательно, пока носители этой интенции апеллируют к нему в его меняющейся физической природе как к тому же самому. Подобным образом мы сохраняем родственные и человеческие отношения с людьми, живые клетки которых полностью обновляются за некий период времени.

Но представим теперь, что в процессе серии реставраций специалисты начали вдруг использовать современные синтетические материалы, созданные с применением инновационных химических и нанотехнологий и компьютерного моделирования. В результате мы получаем изображение, производящее то же самое визуальное и эмоциональное впечатление (сохранены характер и интенсивность оттенков, детали выражения лица, даже текстура холста и красок, видная в увеличительное стекло, – все то, о чем позаботился бы лучший в мире реставратор), но на совершенно иной физической и технологической основе, невозможной в эпоху Леонардо. Можно даже для смелости воображения добавить технологии голографии и 3D-печати. Насколько сократится число людей, ответивших «да» в первом случае?

Поговорим теперь не о живописи, а о фотографии. Произведения этого искусства, имеющие высокую историко-художественную и рыночную ценность, насколько мне известно, могут существовать более чем в одном экземпляре, но количество отпечатков строго ограничено, гарантировано автором или его законными представителями. Любой отпечаток, выходящий за эти количественные пределы, считается в зависимости от обстоятельств обнаружения либо незаконной подделкой, либо легальной копией, т. е. не идентифицируется как то же самое произведение искусства, хотя физически он может ничем не отличаться от зарегистрированных легальных отпечатков. И наоборот, легальные отпечатки могут быть сделаны на бумаге разных типов с применением разных

химикатов, но все легальные, например, двадцать отпечатков – это одна и та же фотография, имеющая примерно одну и ту же аукционную или рыночную цену и распределенная среди двадцати музеев и частных коллекций. И здесь, в отличие от картины Леонардо, нет проблемы ночного выпадения из бытия: в то время как в некотором музее выключен свет и нет посетителей, в другом месте частный коллекционер с любовью рассматривает ту же самую фотографию, радуясь своей удачной инвестиции.

Литературные произведения и их переводы на другие языки представляют еще один особый случай. Здесь физическое воплощение – письменные знаки определенной формы, произносимые при чтении вслух звуки – может быть совершенно различным, но произведение воспринимается как одно и то же.

Зададимся также вопросом, почему, например, авторский пересказ писателем А.М. Волковым «Волшебника страны Оз» Л.Ф. Баума считается отдельным литературным произведением – сказкой «Волшебник Изумрудного города», а авторизованный перевод Б.В. Заходером сказки А. Милна – все тем же «Винни-Пухом». Или почему в музыке в одних случаях исполняемый фрагмент считается законной вариацией известной мелодии, а в других – отдельным произведением, лишь использующим чужую тему? И снова сам собой напрашивается ответ: идентификация произведений искусства зависит от социальных институтов, определяющих в каждой особенной ситуации особенные правила идентификации объектов. В таких правилах качественные аспекты чувственного восприятия могут присутствовать в разной мере и играть разные роли. Иногда достаточно отнесения экземпляра фотографии к множеству зарегистрированных легальных отпечатков, иногда все решает традиция и устоявшееся мнение экспертного сообщества, а бывает, что требуется работа экспертов, которые способны с помощью мелких деталей, заметных лишь профессионально тренированным глазам и мозгу, отличить единственный легальный экземпляр от множества подделок.

Согласно Н. Гудману [Goodman, 1976, p. 113–115], **имеется принципиальное различие между искусствами, по его терминологии, аутографическими (живопись и скульптура) и аллографическими (музыка и литература).** В первом случае подлинным произведением считается один-единственный экземпляр, во втором – произведение создается в виде некоторого сценария, текста или прототипа, который может существовать в потенциально бесконечном множестве различных реализаций – исполнений, постановок, изданий и т. п. Очевидно, что если чувственное восприятие и играет где-то решающую роль в идентификации или оценке произведения искусства, то, конечно, только в случае аутографических искусств. Аллографические произведения, как правило, значительно более схематичны, и качественные особенности чувственных образов там чаще отдаются на откуп производителям реализаций (исполнителям, постановщикам, издателям).

Но вернемся к вопросу, не дающему покоя тем из нас, кто мыслит философски. Почему вообще может возникнуть представление, что картина существует, только пока на нее смотрят, пока длится событие обработки мозгом сигналов, принятых от глаз, и порождения чувственных субъективных – качественных – образов? Потому, будет ответ, что «Моной Лизой» мы считаем данный конкретный образ женщины с темно-русыми волосами, зеленым платьем, определенным овалом лица и перспективой на заднем фоне. И все

это поставляет нам та часть нашего когнитивного аппарата, которая видит мир качественно, в форме субъективных чувственных образов, чьи свойства воспринимаются непосредственно и могут быть только пересказаны, но не переданы другому.

Здесь возникает иная философская проблема: а не видим ли мы разные картины, описывая их одними и теми же словами? Ответ на этот вопрос во многом зависит от того, считаем ли мы, что квалиа – качественные особенности чувственных образов – являются репрезентациями, т. е. представляют какие-либо свойства реальности.

2. Квалиа как репрезентации

Концепция, согласно которой квалиа представляют собой репрезентации, т. е. свойства интенциональных объектов, «представляющие» свойства объектов реальных, может столкнуться со следующей трудностью. «Сильный» репрезентационализм утверждает, что квалиа суть репрезентации и ничего более, природа и назначение их качественной определенности заключается в способности представлять некое содержание, отличное от них самих.

Предположим, что качественный характер слышимых мною тонированных звуков, а именно их высота как чувственно воспринимаемое качество, репрезентирует некоторое свойство реальных объектов – волновых колебаний воздушной среды, а именно их частоту. Когда я слышу такие звуки, упорядоченные определенным образом по законам гармонии, как звучащие одновременно, так и следующие друг за другом, я наслаждаюсь музыкой. Более того, я знаю, что ощущение их гармоничности вызвано действительной математической гармонией частот соответствующих акустических колебаний. И у меня, в самом деле, возникает искушение признать ощущаемую на слух высоту каждой отдельной ноты репрезентацией определенной частоты колебаний воздуха.

Однако я вынужден заметить, что существуют и иные формы репрезентации тех же колебаний – например, их графическое изображение. И, став на минуту «сильным» репрезентационалистом, я должен буду предположить, что зрелище наложенных друг на друга графиков колебаний воздуха должно пробудить во мне такое же воодушевляющее ощущение гармонии, какое я испытываю при прослушивании музыки. Но даже убежденный репрезентационалист, я уверен, признает, что, скорее всего, этого не произойдет. Если даже я увижу некоторую эстетическую ценность в гармонично сочетающихся графиках, это будет наслаждение совсем другого характера, нежели пир звуков, очаровывающий тонкие натуры.

Этот аргумент, как мне представляется, опровергает взгляд, согласно которому квалиа суть *не более чем репрезентации* и объяснение их качественности сводится исключительно к указанию на эту их роль. Мой пример показывает, что независимо от того, являются ли квалиа репрезентациями в принципе, в их качественном своеобразии есть нечто, что обладает каузальной силой, несводимой к их предполагаемой репрезентативной природе, что способно вызывать следствия, недоступные для репрезентаций того же содержания, выполненными средствами другого когнитивного модуля. Поскольку именно не интерпрети-

руемое никакими иными способами качественное своеобразие этих состояний и составляет «трудную проблему сознания», а сильный репрезентационализм оказывается опровергнут моим аргументом, оставшийся «на поле» слабый вариант репрезентационализма не делает нас ближе к ее решению. Ведь если, настаивая на том, что квалиа суть репрезентации, мы тем не менее признаем за ними некий нерепрезентационный качественный «остаток» – в духе концепции «ментальной краски» Г. Хармана [Harman, 1990] и Н. Блока [Block, 1996] – некую часть их качественного своеобразия, которая обнаруживается интроспективно и даже способна самостоятельно участвовать в причинно-следственных отношениях, но в то же время не участвует в репрезентациях, то само указание на способность квалиа транслировать какое бы то ни было интенциональное содержание становится излишним в контексте основной философской проблемы, решение которой мы ищем.

3. Магия комнат и ее разоблачение

По крайней мере в двух из наиболее известных в аналитической философии мысленных экспериментах – «китайская комната» [Searle, 1980] и «комната Мэри» [Jackson, 1986] – имеется существенный элемент лукавства. Конструкция Дж. Сёрла, имеющая целью опровергнуть «сильный искусственный интеллект», строится на том, что условный американец, находящийся внутри изолированной комнаты не понимает китайского, но, следуя правилам, выдает осмысленные ответы на осмысленные вопросы. Вывод: понимание не функционально, а субстанциально: оно зависит от «каузальных сил» конкретной биологической субстанции и, следовательно, не может быть реализовано в другой. Однако, как мне уже доводилось показывать [Михайлов, 2012; Михайлов, 2015, с. 75–76], Сёрл использует иллюзионистский трюк: исподволь переключает наше внимание на человека в комнате, который согласно нашей расхожей картине мира и должен быть субъектом понимания. Тогда как в конструкции эксперимента человек в комнате – всего лишь деталь понимающей системы и не обязан сам по себе обнаруживать понимание. Подобным же образом, отдельный нейрон моего мозга, который в числе прочих возбуждается, пока я пишу эту статью, не расскажет вам о ее целях и содержании.

Но эксперимент Ф. Джексона не менее, а, пожалуй, даже более, лукав. Его цель – опровергнуть физикализм, согласно которому все, что в мире есть нефизического, супервентно по отношению к физическому и полный физический дубликат нашего мира был бы его полным дубликатом [Stoljar, 2017]. Мэри, проведшая полжизни в черно-белой комнате, обладает, согласно Джексону, полным знанием физики, а также физиологии цветовосприятия, поскольку – внимание! – она изучила все имеющиеся теории на эту тему. Тем не менее, выходя из комнаты и впервые видя красочный мир, она, очевидно, получает некоторое новое знание о нем – возражать было бы контринтуитивно. Следовательно, мир не сводится к физическому, в нем есть что-то еще.

Но нетрудно заметить, что логика Джексона строится, по крайней мере, на одной не высказанной явно презумпции: знание о мире изоморфно самому миру. В частности, здесь предполагается, что физический мир адекватно

отражен в физических теориях, а наличие необъяснимых этими теориями феноменов, таких как квалиа, указывает на их принципиально нефизический характер.

Роль квалиа в этом эксперименте может быть двоякой: (1) они полагаются полностью репрезентативными и как таковые демонстрируют знание о чем-то, что не описывается физическими теориями и, следовательно, не относится к физическому миру, или (2) само их наличие опровергает физикализм независимо от их репрезентативности.

Рассмотрим вариант (1). Обширная, инновационная и провокативная (в хорошем смысле слова) аргументация в пользу того, что адекватность перцептивных образов не только сомнительна, но и просто невозможна в силу эволюционных причин, содержится в [Hoffman, 2009]. Но даже наш собственный не столь амбициозный эксперимент с различными репрезентациями звука показал, что чувственные кваликативные феномены или вовсе не являются репрезентациями, или как минимум содержат важную нерепрезентативную часть. Следовательно, их нельзя считать изоморфными миру знаниями о нем. А значит, невозможно утверждать, что, обретя некую новую феноменальность, Мэри узнала что-то о нефизической части мира.

Впрочем, остается возможность, предусмотренная вариантом (2): она не узнала, а непосредственно обрела нечто, что не является физическим. Предполагается, что квалиа сами по себе, вне зависимости от их репрезентационных возможностей, являют собой некую нефизическую часть реальности. И вот здесь мы переходим к главной уловке, скрытой в эксперименте Джексона: к пониманию физического.

Джексон исходит из того, что факт знакомства Мэри со всеми имеющимися на данный момент физическими и физиологическими теориями гарантирует исчерпывающую представленность физической части мира в ее сознании. И тогда любое дополнение к картине – неважно, в виде репрезентации чего-либо неучтенного в ней или в виде его непосредственного присутствия – означает существование нефизического. Можно было бы возразить в очевидном и привычном духе: любая физическая теория исторически ограничена, и, если существование человечества продлится достаточно долго, мы с высокой степенью вероятности узнаем много нового и неожиданного о физическом мире. Но нужно уточнить и то обстоятельство, на которое указывал Блок [Block, 2009, p. 1115]: наиболее существенный прирост научного знания происходит даже не с обнаружением новых фактов, а с изменением концептуальных схем – когда, например, мы вдруг понимаем, что горение, ржавление и метаболизм суть однотипные явления, равно как и свет, звук и круги на воде. В той теоретической картине мира, с которой Мэри была знакома изначально, возможно, отсутствуют концептуальные схемы, которые позволят великим ученым будущего интегрировать квалиа в физическое описание мира.

Таким образом, фокуснический трюк Джексона заключается в том, что своим экспериментом он объясняет неизвестное через неизвестное: трудная проблема сознания состоит в том, что у нас нет хорошей естественнонаучной теории, объясняющей квалиа, и он предлагает эксперимент, которым «доказывает», что физикализм ложен, потому что... у нас нет хорошей естественнонаучной теории, объясняющей квалиа.

4. Еще одна комната

Но даже на том уровне теоретического развития, на котором находимся мы вместе с Мэри, можно прояснить некоторые закономерности взаимодействия чувственного и рационального, феноменального и дискурсивного, чтобы окончательно понять, что же произошло с девушкой.

Представим себе, что в комнате, где мы находимся, на стене сидит муха, а на лампе повисла летучая мышь. В какой-то момент открывается дверь и в комнату входит, например, наш философский оппонент или любой другой человек. Вы слышите звук открывающейся двери; подняв взгляд, вы видите измененное положение двери и некий новый комплекс чувственно воспринимаемых признаков, который соответствует вошедшему человеку. Предположим, что это же событие замечено мухой. Она – существо с фасеточными глазами и совсем по-иному устроенным мозгом. Мы можем попытаться представить себе, каким образом произошедшее событие отразится в ее чувственном мире, но, скорее всего, наша фантазия будет слишком антропоморфной. И совсем не поддается человеческому воображению чувственный образ того же события, созданный эхолотом летучей мыши, героини хрестоматийной статьи Т. Нагеля [Nagel, 1974]. А значит, мы имеем по крайней мере три качественно различных чувственных образа одного и того же события.

Далее, мы ничего не знаем о том, есть ли теоретики-естествоиспытатели среди мух и летучих мышей. Однако среди нас они есть, и физик, химик, физиолог и т. д. опишут произошедшее событие в терминах и в каузальном контексте, которые соответствуют определенной научной дисциплине. В некоторых науках это событие может оказаться нерелевантным и совсем не получить описания. Но точно так же есть события, хорошо описываемые теми или иными научными теориями, но не релевантные нашему перцептивному аппарату, проще говоря, не замечаемые нами.

То, что мы называем научным объяснением, состоит в установлении некоторой взаимосвязи между чувственным образом события и его теоретическим описанием¹. Эта взаимосвязь дает нам новый контекст, из которого становится ясно, почему данное событие произошло именно таким образом, каковы были его причины, каковы последствия, и подобный контекст будет разным в разных научных теориях.

Попробуем разобраться в тех материалах, из которых мы собираем наши чувственные и словесные (или количественные) образы событий. Для создания первых мы используем цвета, «вшитые» в наш когнитивный аппарат, пространственные формы и локализации, также конструируемые известными областями мозга, слуховые ощущения и т. п. Особенности этих «строительных материалов» таковы, что их качественный характер определяется нашим физиологическим устройством и не до конца изученными когнитивными механизмами. В свою очередь поставляемые нашими теориями дискурсивные образы событий состоят из символов, определяемых теми или иными символическими системами с их конвенциональным синтаксисом и конвенциональным семантикой.

¹ Точнее было бы сказать, между обыденным и теоретическим описанием события, представленного также в чувственном восприятии, но не будем усложнять наше рассуждение.

Поэтому, когда мы говорим о физической реальности, доступные способы ее отображения не заимствуют свои «строительные материалы» из нее самой, а обладают ими в определенном смысле а priori. А значит, Мэри, неожиданно увидев мир в цвете, не обрела нового знания, а лишь получила в свое распоряжение некоторые новые изобразительные средства – «строительные материалы» образов реальности, при том что у нее уже имелись качественно отличные образы, относимые научной традицией к той же самой реальности. Как если бы мы вдруг получили возможность чувственно воспринимать движения кварков и бозонов, о которых знаем из физических теорий. Это событие, безусловно, важно как расширение возможностей нашего когнитивного аппарата, но оно ничего не говорит об онтологическом составе реальности.

Репрезентационалистская интерпретация (1) мысленного эксперимента Джексона предполагает, что человек, имеющий непосредственный доступ к цветовосприятию, знает о мире больше, чем человек, никогда не видевший цветов, но знакомый с теорией цветовосприятия. С моей точки зрения, приведенные выше рассуждения демонстрируют недостаточную обоснованность этой интерпретации, поскольку репрезентативной способностью обладают не сами качественно своеобразные элементы чувственного образа, а скорее их наблюдаемые различия, которым ставятся в соответствие символические конструкции, полагаемые репрезентациями тех же фактов и событий. Однако даже если мы встанем на крайне репрезентационалистскую точку зрения и скажем, что качественное своеобразие элементов чувственного образа также представляет собой репрезентацию чего-либо в реальности, то в любом случае из «нефизического» характера репрезентации не следует нефизический характер репрезентируемого. И это при том, что само представление о нефизическом характере средств репрезентации основано на отсутствии объяснения для них в наших текущих физических теориях.

Онтологическая интерпретация (2) истории с Мэри предполагает, что девушке было продемонстрировано нечто, имеющееся в реальности, к чему прежде у нее не было доступа, и это нечто имеет нефизическую природу, поскольку непосредственно не описывается физическими теориями. Однако вспомним на минуту, что у Мэри до переломного события в ее жизни всегда был в распоряжении альтернативный способ описания реальности – с помощью физических теорий. Последние построены из материала символических систем, которые, в свою очередь, также обладают свойствами, не описываемыми физическими теориями: например, семантикой. Если неожиданное открытие квалиа ставит перед Мэри и перед Джексоном онтологические проблемы, то почему их не ставит факт существования самих физических теорий? В чем принципиальная разница между чувственными и символическими средствами отображения с точки зрения их «физичности»? Пожалуй, только в том, что символические системы мы строим сами и целенаправленно, тогда как строительные материалы чувственных образов подарены нам природой.

5. (Не)выразимость и (не)репрезентативность квалиа

Старая философская дискуссия о выразимости квалиа еще более осложняет поиски ответа на вопрос о релевантности репрезентационалистского их понимания. Р. Нат и М. Панда в [Nath, Panda, 2014] пишут, цитируя Витгенштейна, что сказанное об ощущении в третьем лице – это описание, а то, что говорится о нем в первом лице, есть выражение (*expression*). Я полагаю, что все это не настолько строго. На приеме у врача любой из нас может описать свои ощущения в деталях. Более того, детальные описания ощущений с позиции третьего лица достаточно распространены в художественной литературе, поскольку, как представляется, по законам жанра, читатель внутренне отождествляет себя с героем. Иными словами, есть ситуации, в которых имеет смысл описывать – а не просто выражать – ощущения с позиции первого грамматического лица и детализировать внутренний мир кого-то, о ком говорится в третьем лице, как будто он принадлежит говорящему. Следовательно, нет непроходимых логических барьеров между языками описания ощущений в первом и третьем лице.

Можно было бы возразить, что ни в том, ни в другом случае мы не можем указать очевидные условия истинности для этих описаний. Я бы сказал, что оба случая соответствуют правилу формулирования опыта от первого лица: «Если мне кажется, что мне больно, то мне больно». Таким образом, ни пациент, ни литературный персонаж не могут ошибаться в своих ощущениях, потому что, даже если они каким-то образом окажутся неправы, это не окажет никакого влияния на дальнейшее развитие событий, – выражаясь метафизически, не сыграет каузальной роли. А это значит, что никакое репрезентативное содержание – реальное или предполагаемое – не определяет ни природы квалиа, ни их функциональной роли в когнитивных процессах.

Наука, как она развивалась до сих пор, использует язык описания в третьем лице. Следовательно, если требуется научное решение проблемы «невывражимых» феноменальных качественных конститuentов опыта, то философы должны показать, как язык описания феноменального опыта от первого лица может быть переведен в язык описания феноменального опыта от третьего лица.

6. Репрезентации в когнитивной науке

По-видимому, здесь мы достигли пределов философского анализа проблемы репрезентативного характера квалиа, и поэтому есть смысл обратиться к опыту эмпирических наук, в которых репрезентации играют существенную концептуальную роль.

В классической когнитивной науке, как указывает Н. Фреско [Fresco, 2012, p. 356], считается, что репрезентации должны обладать двумя важными свойствами – быть *физически реализуемыми* и быть *интенциональными*. Интенциональность понимается также классически: как наличие значения или содержания, т. е. того, репрезентацией чего она является. Физическая реализуемость предполагает наличие физически допустимого *транспорта* (*vehicle*) репрезентаций, в качестве которых могут выступать вычислительные структуры или состояния мозга. В рамках такого понимания репрезентации действитель-

но являются физически воплощенными сущностями, обладающими семантикой, т. е. символами. Над ними возможны Тьюринг-вычислимые операции, и вся модель когнитивных актов оказывается полностью аналогичной работе компьютера фон-неймановской архитектуры. К несомненным достоинствам классической модели можно отнести ее непосредственную компьютерную реализуемость: один из классиков символического подхода Дж.Р. Андерсон создал компьютерную платформу ACT-R для моделирования когнитивных функций с целью последующей экспериментальной проверки моделей [Anderson, 1983]. Объяснительный принцип построен на научной абдукции: если модель показывает те же результаты, что и живой испытуемый, то с высокой степенью вероятности когнитивный аппарат испытуемого имеет ту же структуру, что и компьютерная модель.

И вот здесь начинаются проблемы. Исторически когнитивистская парадигма торжествует после победоносной полемики Н. Хомского и Б.Ф. Скиннера в конце 1950-х гг. Теория врожденной генеративной грамматики вроде бы смогла объяснить продуктивность человеческого языка – способность составлять и понимать ранее не слышанные высказывания. Лингвистика в союзе с восходящей компьютерной наукой победила необихевиористскую психологию. Неудивительно, что новый когнитивный подход имел ярко выраженные лингвистические родовые черты, опираясь на представление о психической жизни как о потоке вычислительных операций над символами, обладающими семантикой. Закономерным следствием этого видения явилась концепция «мысленного языка» (*Mentalese*), выдвинутая Дж. Фодором. Согласно Фодору, нашим операциям с внешними символами соответствуют внутренние манипуляции с символами-репрезентациями, такие, что символы имеют ярко выраженную знаковую природу, а сами операции аналогичны высказываниям в естественном языке. Именно язык мысли – основа нашего понимания языка общения [Fodor, 1975; Fodor, 2008]. Такая позиция не должна непременно уводить в дурную бесконечность, тем более что Ф. Эттнев еще в 1959 г. представил механистическую модель когнитивного аппарата, позволяющую избежать *парадокса гомункула* за счет перераспределения функций между уровнями организации системы [Attneave, 1961]. Однако, несмотря на преодоление парадокса на уровне дизайна, на концептуальном уровне его опасность сохраняется, если его сформулировать в следующем виде.

(ПГ1:) Чтобы распознать внешнее содержание в символе, необходимы когнитивные способности. Но именно их мы пытаемся объяснить с помощью этой схемы.

Предположим, можно найти техническое объяснение, каким образом когнитивная система распознает синтаксические свойства внутренних символов. Но откуда берется содержание, делающее их интенциональными? И кто это содержание считает? Таким образом, возникает серьезное подозрение, что объясняемое содержится в объяснении. Как замечает Фреско, «внешние репрезентации зависят от внешнего субъекта (*knower*): субъект приписывает объектную семантику структурам данных, строчками и символам» [Fresco 2012, p. 358].

Неслучайно поэтому исследования в рамках символической (классицистской) парадигмы более всего преуспели в объяснении языковых способностей и языковой деятельности.

Кроме того, остается непреодоленной двусмысленность самого понятия репрезентации: имеется ли в виду состояние когнитивного аппарата или психическое (феноменальное) состояние, иными словами, говорим ли мы об *объективных* или *субъективных* репрезентациях? Последние выглядят более законной областью понятия репрезентации, поскольку играют роль представителя объективных положений дел перед мысленным взором субъекта (хотя и это не всегда справедливо). Что же касается объективно регистрируемых состояний когнитивного аппарата, то, на мой взгляд, такое понимание репрезентации играет в классицизме скорее нормативную роль: все исходит из того, что такие репрезентации должны быть, поскольку они предусмотрены в используемых компьютерных моделях.

В рамках различных проектов «brain reading» («чтение мозга») имеются результаты эмпирических исследований, демонстрирующие возможность построения функциональных отношений между паттернами активации определенных отделов мозга и внешними стимулами. Так, известна работа [Pasley et al., 2012], в которой была предпринята попытка с помощью математического моделирования продемонстрировать функциональную связь между произнесенным словом и паттерном активации верхней височной извилины, ответственной за высокоуровневую обработку семантически нагруженной акустической информации. Пациентам, оперируемым в связи с эпилепсией или опухолями мозга, имплантировали в эту область датчики, с помощью которых стало возможно реконструировать структуру нейронных возбуждений, возникающих при восприятии пациентом реальных или специально придуманных слов. Далее было построено несколько математических моделей, описывающих функциональные связи между паттернами активации и волновой формой звучащего слова. Наиболее релевантная модель использовалась для обратной реконструкции акустического образа из нейронных импульсов. Был получен неоднозначный результат: звуковой образ «восстановленных» слов не распознавался слушателями, но визуально тем не менее волновая форма полученного сигнала соответствовала волновой форме реально произносимых человеком слов. Исследователи предположили, что по мере совершенствования технических и математических средств удастся наладить коммуникацию с пациентами, лишенными речи, например, вследствие паралича.

Очевидно, что для конкретной научной области этот результат является промежуточным. Но в рамках концептуального исследования можно предположить, что эмпирический поиск увенчался успехом – найден способ трансляции звукового образа слова в нейронные ансамбли и обратной трансляции. Тогда мы должны принять, что структура ансамбля возбужденных нейронов есть в точном смысле слова объективно регистрируемая репрезентация звука произносимого слова. И это, скорее всего, будет справедливо. Другое дело, что так понятая «репрезентация» не является достаточно операционализированным понятием в контексте когнитивной проблематики и не предоставляет достаточных концептуальных средств для решения философских и когнитивно-психологических проблем, связанных с сознанием и его многочисленными загадками. Тому есть несколько причин.

Во-первых, паттерн нейронного возбуждения в такой же мере оказывается репрезентацией звука произносимого слова, в какой, наоборот, звуковые колебания суть репрезентация структуры возбуждения нейронного ансамбля. А значит, в самом понятии репрезентации, как оно при этом используется, не остается ничего специфически когнитивного.

Во-вторых, подобное расширенное, или «слабое», понимание репрезентации ведет к позиции, которую можно было бы по аналогии с панкомпьютериализмом назвать *панрепрезентационализмом*. Мозговые структуры на тех же основаниях могут считаться репрезентациями внешних событий, на каких синтезированный протеин может считаться репрезентацией цепочки нуклеотидов РНК (или наоборот, что не принципиально).

Таким образом, в основу объяснения мира когнитивных явлений кладется понятие, охватывающее широкий спектр некогнитивных явлений. С точки зрения логики объяснений и определений подобное понятие может в лучшем случае служить родовым, ничего не говоря о видовых отличиях объясняемого феномена. Для теории сознания – философской или психологической – понятие репрезентации в рассмотренной степени разработанности не может быть достаточным.

7. набросок умеренного репрезентационализма

В свете приведенных соображений более рельефно обозначаются методологические недостатки классического репрезентационализма. На мой взгляд, классическое понимание репрезентации исходит из примитивной схемы когнитивного субъекта, окруженного объектами, которые отражаются в представлениях. На самом деле все гораздо сложнее. Вся концепция семантически нагруженных внешних репрезентаций объектов недостаточно обоснована и убедительна. Скорее, можно говорить о субъективных внешних репрезентациях: например, о цвете как репрезентации некоторого спектра электромагнитного излучения. Подобно субсимвольным вычислениям, на которые опирается коннекционизм, было бы правильно говорить о субсимвольных репрезентациях – как внешних, так и внутренних. Например, векторная сумма весов нейронной сети может рассматриваться как представление категориальной структуры данных, на которых сеть была обучена.

Тогда репрезентации имеют смысл только в контексте вычислительных процессов в системах, где такие процессы являются релевантными:

(Def :) Структура А есть репрезентация структуры В в рамках некоторого вычисления, тогда и только тогда, когда в этих же самых рамках А и В связаны устойчивой инвариантной функцией.

Принимая такое «слабое» определение репрезентации, мы оказываемся дальше от конечной цели когнитивной теории, поскольку так понятое репрезентации не обязательно позволяют нам объяснить процесс получения знания в его окончательном виде. Но это единственный способ избавиться от «гомункула» и увидеть познание как процесс в сложной многоуровневой вычислительной системе.

Другая трудность может быть связана с излишне антропоморфным пониманием «семантического» отношения репрезентации к тому, что репрезентируется. Подобно тому, как в человеческом мире знаки и их значения связаны между собой конвенционально, здесь также люди должны *знать*, а следовательно, быть *обучены* этим связям. Если мы позаимствуем эту семантическую теорию для классической версии когнитивной науки, то рискуем столкнуться с другой версией парадокса гомункула:

(ПГ2:) Для того чтобы символическое вычисление было семантически эффективным, когнитивная система должна «знать» семантические отношения между символами и их референтами. Но любое знание представляет собой (или имеет в своей основе) репрезентацию. Тогда каждая репрезентация требует другую, поддерживающую ее репрезентацию и так далее до бесконечности.

В случае чисто синтаксического вычисления мы избегаем этого парадокса, но оставляем механизм, благодаря которому психические состояния вообще имеют содержание, т. е. интенциональность, необъясненным. Но если, потерпев неудачу с компьютерными теориями, мы отступим к чистым и прямолинейным причинно-следственным объяснениям репрезентации, то в конечном итоге упустим смысл всего когнитивного предприятия. Таким образом, это еще один аргумент в пользу более слабого понимания понятия репрезентации, с тем чтобы оно, не будучи необходимым в рамках философии сознания, тем не менее оставалось в когнитивной науке в качестве полезного объяснительного принципа.

Список литературы

- Михайлов, 2012 – Михайлов И.Ф. Искусственный интеллект как аргумент в споре о сознании // *Epistemology & Philosophy of Science / Эпистемология и философия науки*. 2012. № 2. С. 107–122.
- Михайлов, 2015 – Михайлов И.Ф. Человек, сознание, сети. М.: ИФ РАН, 2015. 196 с.
- Anderson, 1983 – Anderson J.R. The architecture of cognition. Cambridge, Massachusetts: Harvard University Press, 1983. 345 p.
- Attneave, 1961 – Attneave F. In defense of homunculi // *Sensory communication: Contributions to the symposium on principles of sensory communication, 19 July – 1 August 1959, Endicott House / Ed. by W.A. Rosenblith*. Cambridge, MA: M.I.T, 1961. P. 777–781.
- Block, 1996 – Block N. Mental Paint and Mental Latex // *Philosophical Issues*. 1996. Vol. 7: Perception. P. 19–49.
- Block, 2009 – Block N. Comparing the Major Theories of Consciousness // *The Cognitive Neurosciences IV / Ed. by M. Gazzaniga*. Cambridge, MA: MIT Press, 2009. P. 1111–1122.
- Fodor, 1975 – Fodor J.A. The Language of Thought. N. Y.: Thomas Y. Crowell Company, 1975. 217 p.
- Fodor, 2008 – Fodor J.A. LOT 2: The Language of Thought Revisited. N. Y.: Oxford University Press, 2008. 228 p.
- Fresco, 2012 – Fresco N. The Explanatory Role of Computation in Cognitive Science // *Minds & Machines*. 2012. No. 22. P. 353–380.
- Goodman, 1976 – Goodman N. Languages of Art: An Approach to a Theory of Symbols. Indianapolis, IN: Hackett, 1976. 277 p.
- Harman, 1990 – Harman G. The Intrinsic Quality of Experience // *Philosophical Perspectives*. 1990. Vol. 4: Action Theory and Philosophy of Mind. P. 31–52.

Hoffman, 2009 – *Hoffman D.* The interface theory of perception: Natural selection drives true perception to swift extinction // Object categorization: Computer and human vision perspectives / Ed. by S. Dickinson, M. Tarr, A. Leonardis, B. Schiele. Cambridge, UK: Cambridge University Press, 2009. P. 148–165.

Jackson, 1986 – *Jackson F.* What Mary Didn't Know // The Journal of Philosophy. 1986. Vol. 83. Iss. 5. P. 291–295.

Nagel, 1974 – *Nagel T.* What is it like to be a bat? // Philosophical Review. 1974. Vol. 4. P. 435–450.

Nath, Panda, 2014 – *Nath R., Panda M.M.* Experience and Expression: The Inner-Outer Conceptions of Mental Phenomena // Indian Philosophical Quarterly. 2014. Vol. I. No. 1–4 (January–December). P. 77–112. (Published in 2015).

Pasley et al., 2012 – *Pasley B.N., David S.V., Mesgarani N. et al.* Reconstructing Speech from Human Auditory Cortex // PLoS Biol. 2012. Vol. 10(1): e1001251. P. 1–13. DOI:10.1371/journal.pbio.1001251.

Searle, 1980 – *Searle J.* Minds, Brains, and Programs // Behavioral and Brain Sciences. 1980. Vol. 3. P. 417–457.

Stoljar, 2017 – *Stoljar D.* Physicalism // The Stanford Encyclopedia of Philosophy (Winter 2017 Edition) / Ed. by E.N. Zalta. URL: <https://plato.stanford.edu/archives/win2017/entries/physicalism/> (дата обращения: 18.06.2018).

Qualia, representations and the magic of philosophical rooms

Igor F. Mikhailov

Institute of Philosophy, Russian Academy of Sciences. 12/1 Goncharnaya Str., Moscow 109240, Russian Federation; e-mail: ifmikhailov@iph.ras.ru

The problem of qualitatively unique and ineffable aspects of the phenomenal experience, known as qualia, forms the basis of the “hard problem of consciousness”, according to popular belief. However, a simple analysis of the role of qualia in the perception and identification of art objects shows that phenomenal images are only one of the building materials of complex cognitions, and their functional role varies from case to case. The paper cites arguments against the anti-physicalist interpretation of qualia, which goes back to the famous “Mary’s room” by Frank Jackson. It shows that neither the epistemological view of this case (qualia as representations of something non-physical in the world), nor the ontological view thereof (qualia as essentially non-physical entities) do not seem sufficiently substantiated when this thought experiment is examined in detail. Further, in the course of conceptual analysis and thought experiments, it is shown that strong representationalism in the philosophy of qualia is irrelevant, while weak representationalism is devoid of the necessary explanatory power. In any case, in the phenomenal subjective images, a non-representational ‘residue’ is present that plays an important causal or functional role in human perceptions. At the same time, moderate representationalism is appropriate and promising in cognitive science, if we consider representations as a necessary element of computing processes, wherever, including in the human cognitive devices. However, it is obvious that, in such a form, this concept does not pretend to be the universal clue for philosophical riddles, and its “import” into philosophy implies the rejection of intentionalist interpretations of representative functions of qualia.

Keywords: qualia, mind, consciousness, representation, representationalism, intentionality, cognitive science, computation

Acknowledgements: The research is carried out at the expense of the Russian Science Foundation grant (Project №16-18-10229).

References

- Anderson, J. R. *The architecture of cognition*. Cambridge, Massachusetts: Harvard University Press, 1983. 345 pp.
- Attneave, F. "In defense of homunculi", in: *Sensory communication: Contributions to the symposium on principles of sensory communication, July 19-Aug. 1, 1959, Endicott House*, ed. by W. A. Rosenblith. Cambridge, MA: M.I.T, 1961, pp. 777–781.
- Block, N. "Mental Paint and Mental Latex", *Philosophical Issues*, Vol. 7, Perception (1996), pp. 19–49.
- Block, N. "Comparing the Major Theories of Consciousness", in: *The Cognitive Neurosciences IV*, ed. by M. Gazzaniga. Cambridge, MA: MIT Press, 2009, pp. 1111–1122.
- Fodor, J. A. *The Language of Thought*. New York: Thomas Y. Crowell Company, 1975. 217 pp.
- Fodor, J. A. *LOT 2: The Language of Thought Revisited*. New York: Oxford University Press, 2008. 228 pp.
- Fresco, N. "The Explanatory Role of Computation in Cognitive Science", *Minds & Machines*, 2012, no. 22, pp. 353–380.
- Goodman, N. *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis, IN: Hackett, 1976. 277 pp.
- Harman, G. "The Intrinsic Quality of Experience", in: *Philosophical Perspectives*, 1990, vol. 4: Action Theory and Philosophy of Mind, pp. 31–52.
- Hoffman, D. "The interface theory of perception: Natural selection drives true perception to swift extinction", in: *Object categorization: Computer and human vision perspectives*, ed. by S. Dickinson, M. Tarr, A. Leonardis, B. Schiele. Cambridge, UK: Cambridge University Press, 2009, pp. 148–165.
- Jackson, F. "What Mary Didn't Know", *The Journal of Philosophy*, 1986, vol. 83, iss. 5, pp. 291–295.
- Mikhailov, I. F. *Chelovek, soznanie, seti* [Man, Mind, Networks]. M.: IF RAN Publ., 2015. 196 pp. (In Russian)
- Mikhailov, I. F. "Iskusstvennyj intellekt kak argument v spore o soznanii" [Artificial Intelligence as an Argument in the Dispute on Mind], *Epistemology & Philosophy of Science / Epistemologiya i filosofiya nauki*, 2012, vol. 32, no. 2, pp. 107–122. (In Russian)
- Nagel, T. "What is it like to be a bat?", *Philosophical Review*, 1974, vol. 4. pp. 435–450.
- Nath, R., Panda, M. M. "Experience and Expression: The Inner-Outer Conceptions of Mental Phenomena", *Indian Philosophical Quarterly*, 2014, vol. I, no. 1–4 (January–December), pp. 77–112. (Published in 2015).
- Pasley, B. N., David, S. V., Mesgarani, N. et al. "Reconstructing Speech from Human Auditory Cortex", *PLoS Biol.*, 2012, vol. 10(1): e1001251, pp. 1–13. DOI:10.1371/journal.pbio.1001251.
- Searle, J. "Minds, Brains, and Programs", in: *Behavioral and Brain Sciences*, 1980, vol. 3, pp. 417–457.
- Stoljar, D. "Physicalism", in: *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), ed. by E. N. Zalta, [<https://plato.stanford.edu/archives/win2017/entries/physicalism/>, accessed on 18.06.2018].