

Междисциплинарные исследования сознания: 30 лет спустя

Е.П. Велихов, А.А. Котов, В.А. Лекторский,
Б.М. Величковский

Аннотация. Данная статья отмечает юбилей публикации Велихова, Зинченко и Лекторского, призвавших в 1988 г. к междисциплинарным усилиям по изучению сознания. Дан общий обзор современного состояния таких исследований в мире. Подробно рассмотрены модели эволюционной многоуровневой организации когнитивно-эмоциональных процессов и поиск мозговых механизмов интермодальной интеграции механизмов восприятия в контексте классической проблемы «sensus communis». В специальном разделе описаны эвристики, используемые для поиска ассоциированной с «Я» информации, а также для обеспечения интеграции и целенаправленного контроля активности мозга в целом. В последнем разделе обсуждаются примеры и ближайшие перспективы практического использования результатов междисциплинарных исследований сознания.

Ключевые слова: сознание, личностный смысл, мозг, дефолтная сеть мозга (ДСМ), уровни когнитивной организации, эволюция, латерализация, гиппокамп, искусственный интеллект, вычислительное моделирование.

Interdisciplinary Consciousness Research: 30 years on

E.P. Velikhov, A.A. Kotov, V.A. Lektorsky,
B.M. Velichkovsky

Abstract. This article celebrates the anniversary of publication by Velikhov, Zinchenko and Lektorsky who argued in 1988 for an interdisciplinary agenda in studying consciousness. At the beginning of article, a short overview of such research world-wide and across disciplines is given. Among main themes considered in more details are multilevel models of evolutionary cognitive-affective organization and studies directed towards discovery of brain mechanisms involved in sensory integration (akin the classical philosophical problem of “sensus communis”) and in processing of personal-sense-related information. In a separate part of the article, two heuristics are described used for an efficient search of Self-referential data and for a goal-directed control of overall brain multilevel architecture. The final part of the article is devoted to an overview of current and emerging technological applications of consciousness research in several important domains of praxis.

Keywords: consciousness, personal sense, brain, Default Mode Network (DMN), levels of cognitive organization, evolution, lateralization, hippocampal formation, artificial intelligence, computational modeling.

Авторы:

Велихов Евгений Павлович — академик РАН, почетный президент Национального исследовательского центра «Курчатовский институт», доктор физико-математических наук, профессор

Котов Артемий Александрович — ведущий научный сотрудник Национального исследовательского центра «Курчатовский институт» и Российского государственного гуманитарного университета, кандидат филологических наук

Лекторский Владислав Александрович — академик РАН, главный научный сотрудник Института философии РАН, председатель международного редакционного совета журнала "Вопросы философии", научный руководитель философского факультета Государственного академического университета гуманитарных наук, доктор философских наук, профессор

Величковский Борис Митрофанович — член-корреспондент РАН, главный научный сотрудник Национального исследовательского центра «Курчатовский институт» и Российского государственного гуманитарного университета, заведующий кафедрой Московского физико-технического университета, доктор психологических наук, профессор

e-mail: velich@applied-cognition.org

Памяти Николая Александровича Бернштейна и
Владимира Петровича Зинченко –
основоположников отечественных исследований
сознания

Ровно 30 лет назад, известные в своих областях знания ученые — физик, психолог и философ — опубликовали в этом журнале статью, посвященную междисциплинарным исследованиям сознания [Велихов, Зинченко, Лекторский, 1988]. Такие исследования, представлявшие тогда делом достаточно далекого будущего, стали за прошедшее время обычной практикой многочисленных научных коллективов в нашей стране и за рубежом. По сравнению с тем ранним манифестом основные изменения затронули три области. Во-первых, исследования сознания стали опираться на естественнонаучные методы, связанные, прежде всего, с изучением работы мозга. Во-вторых, их существенным компонентом стали задачи вычислительного моделирования, в перспективе ведущие к созданию виртуальных агентов и роботов, в той или иной мере демонстрирующих свойства сознания в своем поведении. Наконец, в-третьих, они стали во многом носить прикладной характер. Мы остановимся на этих аспектах исследований более подробно после краткого обзора панорамы работ в мире.

Методологический плюрализм и новая монадология

Современная литература по проблематике сознания чрезвычайно обширна. Библиография основных источников может быть найдена в ряде работ [Величковский 2015; Лекторский 2010; Bridgeman 2015; Fox, Christoff (eds) 2018; Liu, Perry (eds) 2014]. Вторжение естественнонаучных подходов и инструментов привело к потоку эмпирических исследований, в которых использование новых технологий эксперимента (нейровизуализация, айтрекинг, виртуальная и расширенная реальность) сочетается с опорой на интроспективные отчеты. Направленность большинства таких работ на отдельные, узко локализованные механизмы обнаруживает влияния концепции «модулярности познания» (см. [Lectorsky 2017; Velichkovsky 2015]). В этой мозаичной, быстро усложняющейся картине размываются контуры классических когнитивных функций. Так, уже давно обсуждаются принципиальные различия нескольких систем внимания и восприятия в отношении эволюционного происхождения, мозговых механизмов и вовлеченности эффектов сознания. Известно несколько различных систем памяти, а в работах нобелевского лауреата Д. Канемана мышление оказалось расщепленным на два вида [Канеман 2017]. Эти процессы затронули и понимание сознания, которое описывается как состоящее из нескольких форм и видов, чему способствует несовпадение значения соответствующих терминов в различных языках.

В когнитивной психологии сознание часто идентифицируется с вниманием и рабочей памятью, структура которой включает, однако, ряд служебных механизмов, к сознанию, заведомо не относящихся. Не совпадает сознание и со вниманием, так как последнее обеспечивает направленность психических процессов на решение некоторой конкретной задачи. Понятие «сознание» охватывает более широкий круг феноменов: мы можем интроспективно наблюдать типичный поток сменяющихся друг друга впечатлений и мыслей

в состоянии покоя, при отсутствии какой-либо явной задачи или же когда «задача» имеет чрезвычайно общий характер, например, размышления о смысле жизни.

Несколько эклектичным объединением этих и более ранних представлений является теория глобального рабочего пространства Бернарда Баарса [Baars 1997]. Он описывает сознание, используя известные метафоры «верстака» и «театральной сцены». На сцене и за сценой выполняется множество когнитивных операций, а сознание является как бы лучом прожектора, выхватывающим из рабочей памяти то, что привлекло внимание. Сцена предоставляет возможность для взаимодействия различных сетей мозга, конкурирующих за решение возникающих задач, позволяя оперировать с семантикой человеческой речи, учиться, контролировать действия и, как следствие, формировать понятие "Я" как автора выполняемых действий. Аллюзии на метафору "сцены" можно найти в посвященных сознанию работах биологов, например, в теории повторного ввода (re-entry) нобелевского лауреата Дж. Эдельмана [Edelman 1990].

Среди философских подходов выделяются концепции Дэниела Деннета и Дэвида Розенталя, вышедшие на этап эмпирической проверки. Деннет пытается исключить интроспективные категории из теории сознания. Многолетняя борьба с гомункулусом, внутренним наблюдателем событий на сцене «картезианского театра», привела его к анализу эволюционного развития мозга и к теории сознания, описываемого как облако вероятностей возможных восприятий. Подобно тому, как в квантовой физике, параметры элементарных частиц остаются неизвестными до момента измерения, акт осознания, согласно Деннету, реализует лишь один из множества «проектов» развития событий¹. Теория «множественных проектов» иллюстрируется феноменами стробоскопического движения [Dennett 1992]. Работы Розенталя, первоначально возникшие вне междисциплинарного контекста, подчеркивают способность человека к рефлексии и самосознанию [Rosenthal 2005]. Необходимым признаком сознательного состояния является наша осведомленность об этом состоянии. Такая осведомленность не дана просто в силу активации областей мозга, а достигается за счет "мыслей более высокого порядка" (higher-order thoughts, HOTs), объектами которых становятся первичные, не осознаваемые состояния, восприятия и мысли. В последнее время проводится проверка нейропсихологических следствий этой теории в отношении сенсорных качеств и оценок уверенности в правильности воспоминаний [Lau, Rosenthal 2011; Rosenthal 2018].

Нейрофизиологическая линия, безусловно, является доминирующей в современных исследованиях. Никогда ранее в истории изучения сознания – до и после Декарта – анатомия и физиология мозга не занимали в ней столь важного места. Так понимаются сегодня и когнитивные науки, к названию которых часто даже не считается нужным добавлять «нейро» – связь познания и сознания с мозгом представляется очевидной. В нейрофизиологии акцент делается на изучении мозга животных, прежде всего мышей и крыс. В контексте нашей статьи такая стратегия проблематична – не только потому, что

¹ Деннет, однако, явно дистанцируется от квантово-механических трактовок сознания, особенно в вопросе свободы воли. Он отмечает, что неопределенность микромира никак не помогает личности контролировать свои поступки, а равно нести за них моральную ответственность [Dennett 2003]. В отечественной литературе также анализируется связь квантовых эффектов и феноменов сознания [Петренко, Супрун 2017; Соколов 2010].

сознание у животных все еще остается гипотезой, но и из-за отличий мозга животных и человека: клеточная и субклеточная организация эволюционно наиболее новой фронтальной коры мозга человека заметно отличается от таковой у грызунов и даже приматов [He et al. 2017]. Вместе с тем, роль морфологии мозга может быть сильно преувеличена: птицы, некоторые виды которых решают тесты на рефлексивное сознание успешнее приматов, не имеют коры. Их достижения связаны с субкортикальными механизмами базальных ганглиев и архикортексом, в частности, гиппокампом. Первичной функцией гиппокампа у птиц [Striedter 2016] и млекопитающих [Moser E., Moser M. 2008] является пространственная память, предполагающая элементарные формы ориентации в пространстве.

Для рассматриваемой проблемы особенно важны две группы методов – визуализация макромасштабных нейросетей головного мозга человека и молекулярно-генетические исследования. При всей важности молекулярно-генетических работ, они пока еще находятся в самой начальной стадии развития. Так, недоказанным остается предположение, что сознание может основываться на квантовых эффектах функционирования белка тубулина в микротрубочках нейронов [Соколов 2010; Hameroff, Penrose 2014], а сами методы молекулярного анализа могут быть использованы лишь по отношению к постмортальным образцам тканей. Центральное положение в методическом арсенале нейрокогнитивных исследований поэтому занимают полностью неинвазивные методы визуализации работы мозга, прежде всего, функциональная магниторезонансная томография (фМРТ) и электрофизиологические методы, электроэнцефалография (ЭЭГ) и магнитоэнцефалография (МЭГ). Для проверки возникающих в результате нейровизуализации гипотез о причинно-следственных связях мозговых механизмов эти методы должны быть дополнены возможностями прямого физического воздействия, такими как транскраниальная магнитная стимуляция (ТМС) и стимуляция мозга слабым электрическим током. В последние годы появились и чисто математические методы проверки таких гипотез, прежде всего метод динамического каузального моделирования, использующий данные о временных рядах электромагнитных колебаний и/или флуктуаций метаболической активности структур мозга [Sharaev et al. 2016]. В отличие от электрофизиологических измерений, в случае фМРТ речь идет об относительно низкочастотных колебаниях, сопоставимых в отношении временных характеристик с динамикой ментальных процессов по данным интроспекции.

Вычислительное моделирование важно не только в контексте обработки нейрофизиологических данных, но и как самостоятельный подход к изучению сознания. Среди специалистов по компьютерным наукам, распространена точка зрения, что сознание можно моделировать в виде архитектуры обработки данных ещё до того, как будет открыта физическая природа механизмов, лежащих в его основе [Chalmers 1996]. Какую бы основу не имело сознание – нейросетевые взаимодействия, отношения идеальных сущностей в сфере семантики или квантовые процессы, его можно представить как когнитивную архитектуру – механизм обработки данных. Если реализация функций сознания представляет собой аналог известных нам алгоритмов, то задача моделирования этих функций становится практически достижимой.

По мнению Марвина Минского, для «самосознания» требуется, чтобы некоторый когнитивный механизм человека получал доступ к другим когнитивным механизмам, моделировал и видоизменял их [Minsky 1988]. Предположим, что человек (обозначим его М) обладает моделью себя М*, которая помогает ответить на вопросы о физических признаках, а также моделью более высокого уровня М**, которая содержит описания психологических качеств М* и других людей. Поскольку физические и ментальные характеристики представлены у человека на различных уровнях моделирования, противопоставление М* и М** ведёт, как считает Минский, к психофизиологической проблеме. В современных исследованиях подобный подход реализован в системе «двухуровневой интроспекции» [Valitutti, Trautteur 2017] и модели CogAff [Sloman, Chrisley 2003], разработанных для управления поведением виртуальных агентов и робототехнических систем. При всей привлекательности используемых аналогий правдоподобность вычислительных моделей сознания зависит от учета эволюционных предпосылок возникновения уровневых архитектур.

Уровни когнитивно-эмоциональной организации

Заметный вклад в эти работы могут внести идеи более раннего периода развития науки, в частности, разработанные основателем современной биомеханики Н.А. Бернштейном [Бернштейн 1947] представления об эволюционной организации процессов построения движений человека. В своих работах он детально описал четыре уровня такой организации, от относительно примитивных процессов палеокинетических регуляций и синергий до целенаправленных движений уровней «пространственного поля» и предметных действий. Исследования последних десятилетий во многом подтвердили эти предположения, например, в отношении эволюционной диссоциации механизмов движений «пространственного поля» и движений, учитывающих особенности формы объектов. Такое же разделение обнаружено сегодня в механизмах регуляции движений глаз и в процессах восприятия, где эволюционно первичной является функция пространственной локализации [Velichkovsky et al. 2005]. На базе быстрой и, как правило, не требующей сознательного контроля пространственной локализации осуществляются идентификация и категоризация объектов. Наше восприятие, таким образом, разделяет вопросы «Где?» и «Что?», причем ответом на эти вопросы заняты специализированные нейросети: дорсальный и, соответственно, вентральный потоки переработки информации [Milner, Goodale 2008].

Как известно, Н.А. Бернштейн противопоставлял развивавшуюся им «физиологию активности» стимульно-реактивным схемам. Современной демонстрацией принципа активности является повышение энергопотребления и согласованности флуктуаций в группе центральных структур головного мозга человека как раз при полном отсутствии внешней стимуляции — в состоянии бодрствующего покоя, представляющем собой базовое состояние нашего сознания [Величковский, Ковальчук, Ушаков, Шараев 2016; Knyazev et al. 2017; Raichle 2015]. Эти структуры были названы «дефолтной сетью мозга» (ДСМ, от англ. Default Mode Network, DMN). Они включают медианные отделы лобных долей, оба гиппокампа и несколько структур задних отделов новой коры, в частности, височно-теменно-затылочные области левого и правого полушарий, обеспечивающие интеграцию сведений от основных сенсорных систем — зрения, слуха, обоняния,

соматосенсорной чувствительности и вестибулярного аппарата [Sharaev et al. 2016]. Накапливаются данные, позволяющие считать ДСМ в целом ответственной за поддержание когнитивно-эмоциональных процессов во внутреннем мире человека, автобиографические воспоминания, планирование будущего и интроспекцию [Raichle 2015]. С активностью ДСМ отрицательно коррелирует работа обширной группы механизмов, связанных с решением конкретных задач. Эти механизмы получили название «сетей, связанных с задачей», или «антидефолтных сетей мозга» [Di Perri et al 2016].

Очевидно, редкие периоды покоя используются нами для выявления проблем, имеющих особую личностную значимость, и внутренней работы над ними. Как когда-то отмечал Бергсон, присутствие задачи может исказить работу сознания, сужая его до образа цели действия. Новые экспериментальные результаты говорят о правополушарном доминировании тех структур дефолтной сети и ее окружения в новой коре, которые связаны с «личностной образующей» сознания [Velichkovsky et al. 2018]. В.П. Зинченко, вслед за А.Н. Леонтьевым, называл этот аспект сознания «личностным смыслом». Существование таких механизмов Н.А. Бернштейн предвосхитил буквально в одной строчке своего труда 1947 года, предположив, что «один или два» уровня «высших символических координат» расположены над выделенными им уровнями построения движений. Действительно, предпринятая недавно попытка семантического картирования категорий естественного языка [Huth et al. 2016] на поверхность коры головного мозга показала, что семантические карты языка довольно точно повторяют очертания ДСМ. К удивлению авторов при этом не наблюдалось преимущественной локализации семантических категорий в левом полушарии, со второй половины 19-го века считающегося собственно «речевым» и «доминантным» в отношении произвольной активности (см. [Лурия 1969]). Более того, семантические категории, имеющие оттенок личностной, социальной и эмоциональной значимости чаще отображались на передние и задние отделы коры правого полушария.

Аналогичное развитие имело место в связи с изучением другой системы мозга, часто упоминаемой как один из нейрофизиологических механизмов сознания, а именно, системы зеркальных нейронов, локализованной в нижних отделах префронтальной и теменной коры. Зеркальная система мозга (ЗСМ) была открыта в экспериментах на макаках нейрофизиологами из университета Пармы, Джакомо Риццолатти и его сотрудниками. Они обнаружили, что примерно 10% нейронов в этих отделах мозга одинаково реагирует на собственные действия животного и на такие же действия, выполняемые другим животным или человеком. В начале этого века использование функциональной нейровизуализации позволило обнаружить аналогичные механизмы в передних и задних отделах коры головного мозга человека. Наряду с гипотезами о роли ЗСМ в имитационном обучении и в эффектах интерсубъективности, распространенным является мнение, что она служит основой для формирования механизмов речи. В этом случае в работе ЗСМ следовало бы ожидать левополушарную асимметрию. Экспериментальные данные говорят об обратном – локус максимальной активности в соответствующих задачах был обнаружен в височно-теменной области правого полушария [Aziz-Zadeh et al 2006].

Метаморфозы «sensus communis»

Сквозной темой в размышлениях об интегративной функции души и сознания является предположение о возможном обеспечении такой интеграции за счет представленности в одном из органов тела всех видов чувствительности, без выделения отдельных модальностей, а также левых или правых областей окружения. По Аристотелю, местом такого объединения чувств, «*koine aesthesis*», служит сердце: оно является непарным органом и явно реагирует на существенные «движения души». В переводе на латинский, приписываемом Фоме Аквинскому [Lisska 2016], «*koine aesthesis*» стало звучать как «*sensus communis*», получив в истории философии дополнительно к перцептивному еще и рассудочное значение, предельно выраженное в английском переводе: «*common sense*» («здравый смысл»). Тем не менее, именно аристотелевское понимание использовалось Декартом². К началу 17-го века стала понятной роль сердца как насоса, перекачивающего кровь, и поиски органа «объединения чувств» сместились на мозг. О результатах анатомических опытов, приведших к описанию шишковидной железы (*corpus pineale*), Декарт пишет в канун Рождества 1640 г. своему другу и наставнику Маре Мерсьенну: «И поскольку это единственная непарная часть всего мозга, она с необходимостью и есть место *sensus communis*, иначе говоря, мысли, а, следовательно, и души, ибо одно не может быть отделено от другого» [Descartes 1640].

В последние десятилетия, эта точка зрения упоминалась главным образом как исторический курьез, одна из «ошибок Декарта». Известно, впрочем, что шишковидная железа участвует в выработке гормона мелатонина, регулирующего режим сна и бодрствования и, таким образом, вносит вклад в базовые изменения сознания. Основной критический аргумент связан не с анатомией (отмечают, например, симметричность внутренней структуры этого органа), а с эволюцией. Шишковидная железа представляет собой часть среднего мозга, эволюционно соответствующего примерно стадии появления земноводных. Возможные формы чувствительности ограничены в этом случае острой болью и контрастами света и тени [Lacalli 2018]. О рефлексивном сознании, которое имел в виду Декарт, конечно, не может быть и речи. Кроме того, уникальность органа («отсутствие симметричной пары») должна рассматриваться в отношении функций. Так, левое полушарие коры в анатомическом отношении примерно симметрично правому³, но уникально у большинства людей в отношении построения грамматически правильных речевых конструкций.

Вместе с тем, нельзя объявить задачу поиска «*sensus communis*» просто псевдопроблемой. Дело не столько в авторитете Аристотеля и Декарта, которые эту задачу поставили, сколько в существовании множества клинических фактов, говорящих, во-первых, о хрупкости механизмов сознания и, во-вторых, их выраженной латерализации в координатах мозга. Грубые нарушения сознания – от игнорирования левой половины окружающего пространства, до проблем, связанных с самооценкой, а также с оценкой знаний, намерений и эмоций других людей – чаще всего наблюдаются при поражениях правого полушария.

² К Аристотелю при обосновании эстетики возвращается и Кант.

³ Здесь мы несколько упрощаем реальное положение дел, так как известны систематические нарушения симметричности левого и правого полушарий [Dolina et al. 2017].

Данные о роли правого полушария в клинических эффектах получают объяснение в экспериментах на здоровых испытуемых, упомянутых в предыдущем разделе. В них была предпринята попытка определить взаимоотношения структур внутри и между основными эволюционными слоями коры головного мозга человека в базовом для сознания человека состоянии бодрствующего покоя [Velichkovsky et al. 2018]. На фоне доминирования эволюционно более новых структур контрастно выделялась активность одной из структур древнейшей коры, а именно правого гиппокампа. Его влияние прослеживается по всей «эволюционной вертикали»: от древнейшей к древней и к новой коре, включая наиболее новые фронтально-полярные области. Дополнительный анализ показал, что левый гиппокамп получает информацию только от ипсилатерального (левого) полушария коры и поэтому «видит» лишь противоположную, правую половину собственного тела и окружения. Правый гиппокамп, напротив, получает интермодальную сенсорную информацию как от левого, так и от правого полушария, т.е. в этом отношении вполне может претендовать на статус органа «sensus communis».

Еще одним кандидатом для роли «центра сознания» является другая субкортикальная структура мозга человека – клауструм («ограда»). Эта парная структура расположена вблизи гиппокампов и тесно с ними связана. В статье нобелевского лауреата Френсиса Крика и Кристофа Коха [Crick, Koch 2005] гипотеза о роли «ограды» в интеграции модальностей была обоснована данными о чрезвычайно широкой и плотной системе ее связей с другими областями мозга. Последующие трактографические работы подтвердили уникальную плотность связей. В клиническом контексте был также описан случай обратимого «отключения» сознания пациента путем электрической стимуляции клауструма. Вместе с тем, исследования не смогли выявить признаков интермодальной интеграции: ответы клауструма на сенсорную стимуляцию оказались мономодальными [Remedios, Logothetis, Kayser 2010]. Нет данных и о структурной или функциональной латерализации этой структуры, а сравнительные исследования не обнаружили каких-либо значимых изменений ее размеров при переходе к человеку (увеличение было обнаружено у дельфинов – см. [Baizer et al. 2014]). По-видимому, речь идет скорее о релейном центре, функции которого следует рассматривать в связи с работой гиппокампальной формации.

Сознание и эвристики работы мозга

Хотя решением психофизиологической проблемы в течение многих лет с переменным успехом занимаются представители сразу нескольких научных дисциплин, наш мозг каким-то образом легко переформатирует свою активность при возникновении новых целей, безошибочно различая сведения, связанные с физическими и ментальными признаками [Vogeley 2017]. Складывается впечатление, что при этом используются две эвристики. Первая эвристика позволяет относительно быстро и эффективно найти в нейросетевой активности информацию, связанную с различными аспектами собственного и чужого «Я». Вторая делает возможным целенаправленное функционирование мозга в целом, несмотря на колоссальную сложность его многоуровневой сетевой архитектуры.

В нейрпсихологии с конца 80-ых годов прошлого века высказывались предположения о связи правых отделов лобной коры с рефлексивным сознанием. Исследования последних

лет обнаружили аналогичную связь и с задними отделами коры правого полушария, в области пересечения третичных полей височных и теменных долей [Velichkovsky et al. 2017]. Но если для проявлений сознания важны и передние и задние отделы новой коры, то между ними должно быть какое-то критически важное связующее звено. Судя по всему, это правый гиппокамп, точнее, парагиппокампальная извилина, в которую в результате проекции дорсального потока поступает информация о пространственном окружении [Dickerson, Eichenbaum 2010]. Хотя речь идет о древнейшей коре и специфические для нее формы представления информации относительно элементарны, в них уже присутствует "Я". Первоначально оно присутствует неявно, как центральная зона *эгоцентрической* репрезентации окружения⁴. Но в дальнейшем через этот интерфейс может осуществляться обработка любой ассоциированной с "эго/Я" информации — от пространственной и телесной до эмоциональной и когнитивной; на самом деле, вплоть до декартовского "Я мыслю, следовательно, я существую (существует мое мыслящее «Я»)!"

Гиппокампальная формация левого полушария не имеет доступа к целостной репрезентации пространства и поэтому не может воспользоваться этой простой эвристикой доступа к релевантным «Я» знаниям. Судя по разрозненным литературным данным, она специализируется на долговременной, прежде всего вербальной памяти. Однако, если задача затрагивает «эго/Я» человека — автобиографические воспоминания, кодирование информации в терминах личностного смысла, планирование своего будущего с учетом интересов и намерений других людей, существенной становится работа правополушарных механизмов [Habib, Nyberg, Tulving 2003], исходно кодирующих пространственную информацию.

Второй эвристикой является использование сознания для гибкого изменения и уменьшения сложности функциональной архитектуры мозговых процессов. В референтной для нас статье [Велихов, Зинченко, Лекторский, 1988] подробно обсуждались бытийный и рефлексивный аспекты сознания, к которым В.П. Зинченко добавил в своих поздних публикациях [Зинченко 2010] еще и духовный. Но обычно разделение остается бинарным, что было заметно уже в исследованиях гештальтпсихологов, отмечавших обязательность разделения феноменального поля на фигуру и фон. Концептуально более глубокое различие можно найти у Н.А. Бернштейна [Бернштейн 1947]. Он считал, что в иерархии уровней построения движений нами осознается «смысловое содержание» целей ведущего уровня, т.е. той группы механизмов мозга, операционные ресурсы которой наилучшим образом соответствуют особенностям актуальной ситуации. Остальные уровни работают в фоновом режиме, при котором операции не осознаются. Если ситуация меняется, то любой из уровней может получить статус ведущего, а реализуемые им цели начать осознаваться, так что на их достижение потенциально могут быть направлены совокупные ресурсы организма.

В настоящее время работа мозга и феномены сознания чаще всего рассматриваются как процессы вычислений [Dehaene, Lau, Kouider 2017]. При значительном числе эволюционных уровней когнитивно-аффективной организации и огромном количестве

⁴ Наглядным примером в рамках одной модальности может быть тень от носа в центре «циклопического» поля зрения.

входящих в них компонентов, сходимость подобных вычислений не может быть гарантирована. Система была бы вынуждена работать как набор автономных модулей. Этот *modus operandi* естественен и единственно возможен для технических устройств, но у человека он наблюдается только при нарушениях сознания или в лабораторных условиях, когда испытуемые не знают о подлинных целях эксперимента [Velichkovsky 2002]. Бинарное разделение множества механизмов на ведущий уровень, цели которого осознаются, и «все остальное» (фоновые уровни) радикально упрощает задачу координации работы мозга. В этом отношении можно говорить о сознании как виртуальной машине, гибко меняющей свое положение на эволюционной шкале механизмов мозга, а также о «минимальной архитектуре» сознания, состоящей всего лишь из двух уровней [Kotov 2017].

Еще не совсем понятно, как эта эвристика имплементируется с помощью известных механизмов мозга. Наряду с рассмотренными выше процессами, важную роль могут играть функции фронтальной коры [Burgess, Cohen-Yaacovi, Volle 2012; Penfield, Evans 1935], позволяющие сопоставлять объективированные и подчеркнута субъективные репрезентации одного и того же содержания. Это является необходимым шагом в социальных взаимодействиях, при рассуждениях в модальности «как если бы», а также при понимании (порождении) метафорических сравнений, юмора и иронии [Kotov 2017; Minsky 1988]. Трактовка сознания как механизма работы с множеством репрезентаций отдаленно напоминает теорию «множественных проектов» Деннета, ориентированную на объяснение феноменов восприятия [Dennett 1992]. Наша более широкая трактовка открывает путь для моделирования воображаемых ситуаций, собственных возможных действий в них и возможных действий других агентов, т.е. обнаруживает признаки, считающиеся существенными признаками рефлексии в философии познания и когнитивной психологии.

Исследование сознания как прикладная проблема

За прошедшие 30 лет в исследованиях сознания произошло много важных изменений, но, пожалуй, главным является то, что они стали практически значимыми. Это относится и к сенсорному и к действенному аспектам сознания, т.е. к изучению, с одной стороны, «феноменальной ясности» осознаваемых содержаний и, с другой, субъективной свободы выбора принимаемых решений [Лекторский 2017]. Под заметным влиянием работ В.П. Зинченко субъективное переживание вышло на первый план в эргономике, тогда как в экономике рыночный успех и эффективность финансовых инвестиций оказались в значительной степени зависящими от психологических aberrаций сознания, «когнитивных иллюзий» [Канеман 2017]. Индивидуальное и социальное сознание находятся в фокусе политтехнологий, использующих всю мощь современных средств информационного воздействия.

Нейрокогнитивная революции не обошла стороной и эти области, в результате чего появились новые прикладные дисциплины: нейроэргономика, нейроэкономика и нейромаркетинг. Некоторые из нейрокогнитивных и поведенческих методов развиваются в направлении все более эффективных технологий экспликации индивидуальных знаний и содержания сознания [Величковский 2017]. «Внутренний театр для себя» (выражение

Н.Н. Евреинова) часто включает нескольких действующих лиц: мы замечаем, что ведём с собой или кем-то другим внутренний диалог, смотрим на себя со стороны глазами окружающих, оцениваем других в зависимости от того, как они оценивают нас, пытаемся представить, как бы мы поступили на месте другого либо как другой повел бы себя в нашей ситуации [Петровский 2013]. Интерсубъективность сознания изучается в социальных нейронауках и том их разделе, который направлен на кросскультурные исследования (cultural neuroscience).

Диагностика сознания всегда играла важнейшую роль в медицине. Коммуникативный контакт с пациентом остается основным тестом сохранности сознания. Одновременно растает значение объективных методов, объединяющих нейронауки и математические модели. Примером служит реализация теории интегративной информации [Tononi, 2012], в которой сознание рассматривается как «единая, субъективно самоощущаемая субстанция». Данный взгляд на сознание был выражен с помощью количественного коэффициента, оценивающего комбинаторную сложность ответов мозга на ТМС. Использование этого коэффициента позволяет различать пациентов, находящихся в состояниях ясного сознания, сна с движениями глаз, сна без движений глаз, наркоза и комы [Casali et al., 2013]. К числу преимущественно медицинских приложений относятся также интерфейсы «мозг-компьютер» и «глаз-мозг-компьютер» [Величковский и др. 2016; Shishkin et al. 2016]. Они позволяют находить признаки произвольного намерения в сигналах ЭЭГ/МЭГ и движениях глаз, преобразуя интенцию в движения робототехнических устройств. Решение этой проблемы завещано нам Декартом: «Как мысль о руке поднимает руку?». Иронично, что именно эта проблема, практически решаемая в наше время методами айтрекинга и электрофизиологии, была когда-то объявлена основателем электрофизиологии Эмилем Дюбуа-Реймоном относящейся к категории «Не знаем и никогда не узнаем».

Прорывной областью развития технологий сегодня становится машинное обучение. На базе использования алгоритмов «глубокого обучения» искусственных нейросетей с промежуточными слоями и высокопроизводительных графических процессоров удалось на приемлемом уровне решить такие десятилетиями не поддававшиеся решению задачи, как распознавание речи, компьютерное зрение и машинный перевод. В играх с фиксированным набором правил машины уже сейчас демонстрируют сверхчеловеческие способности, которые быстро прогрессируют с ростом компьютерных мощностей. Несмотря на впечатляющие успехи, у этого поколения программных продуктов все еще отсутствуют характерные для человеческого интеллекта гибкость и способность работать в новых условиях. Так, хотя программа AlphaGO и позволила недавно обыграть мирового чемпиона по игре Го, для этого ей понадобился просмотр порядка 100 млн. игровых ситуаций, тогда как сам чемпион, по-видимому, опирался на опыт менее чем 50 тысяч партий [Lake et al. 2017]. Различия имеют качественный характер. Игрок средней квалификации способен начать играть по *ad hoc* правилам и в необычных условиях – с доской другого размера или формы, например, представляющей собой поверхность ленты Мёбиуса. Эти изменения блокируют «машинный интеллект», причем преодолеть блокировку можно лишь путем усилий высококлассных программистов и новой фазы обучения. Гибкость и способность работать в новых условиях относятся к конституирующим свойствам сознания, определяющим силу человеческого мышления.

Моделирование сознания и соответствующая имплементация представляют собой, таким образом, важнейшее условие дальнейшего развития работ в области когнитивных технологий и «нейроинтеллекта».

Аналогичный вывод можно сделать и в связи с развитием речевых технологий, где использование нейронных сетей пока наталкивается на непреодолимые трудности, как только требуется художественный перевод или же диалог ведется в режиме свободного общения. Задача автоматической обработки текста состоит в том, чтобы построить одно – правильное – понимание, а задача системы поддержания диалога – предложить пользователю наиболее релевантный ответ. Человек, очевидно, действует иначе: даже в серьезном диалоге мы можем шутить или предлагать ироничный ответ, намеренно используя смысловую многозначность и речевую омонимию. Дальнейшее развитие этой прикладной области потребует серьезной работы с логико-философскими вопросами лингвистики, включая анализ особой роли психологических глаголов и маркеров «ментальных пространств» в построении образа ситуации [Fauconnier 2018]. Иными словами, там, где это не случилось ранее, проблема сознания начинает занимать центральное положение как крупнейшая из еще нерешенных проблем науки.

Заключение

Прошедшие три десятилетия оправдали позитивный настрой первой в отечественной науке работы, призвавшей к междисциплинарному изучению сознания [Велихов, Зинченко, Лекторский, 1988]. Трудно было предположить тогда, каким мощным окажется потенциал подобных исследований. Некоторые из полученных результатов, новые методы и новые проблемы рассмотрены нами в этом обзоре с общих позиций методологического реализма [Лекторский, 2017]. Пристальное внимание к материальным основам сознания отнюдь не свидетельствует о приверженности редукционизму в варианте теории идентичности мозга и сознания. Тектонические сдвиги в понимании сознания, при которых мы присутствуем и в которых участвуем сегодня — это то, что происходит раз в два-три столетия, своего рода поздний триумф натурфилософии. Идет ли речь о нейрофизиологических, молекулярно-генетических или вычислительных моделях, сложность, разнообразие и непрямой характер современных методов исследования лишь расширяют поле возможностей для поиска строго научных объяснений и интерпретации результатов в рамках социогуманитарных и философских подходов.

Благодарность

Подготовка этой статьи частично поддержана грантами РНФ 17-78-30029 (картирование семантических категорий русского языка) и 18-19-00593 (человеко-машинное взаимодействие с использованием маркеров ожидания и намерения), грантами РФФИ офи_м 16-29-09601 (парсер для выявления эмоциональных и экстремистских суждений), офи-м 17-29-07083 (моделирование социального поведения автономных роботов), внутренним грантом НИЦ «Курчатовский институт» (приказ №1378 от 23.08.2017), направленным на изучение уровневой организации мозга человека в интерфейсах мозг-компьютер, и субсидией Министерства образования и науки РФ №14.601.21.0017 (Уникальный идентификатор работ RFMEFI60117X0017).

Литература

Aziz-Zadeh et al 2006 — Aziz-Zadeh L., Koski L., Zaidel E., et al. (2006). Lateralization of the human mirror neuron system. *Journal of Neuroscience*, 26(11), pp. 2964-2970.

Baars 1997 – Baars B.J. (1997). *In the theatre of consciousness*. NY: Oxford University Press.

Baizer et al. 2014 – Baizer J. S., Sherwood C. C., Noonan M., Hof P. R. (2014). Comparative organization of the claustrum: what does structure tell us about function? *Frontiers in Systems Neuroscience*, 8: 117.

Бернштейн 1947 — *Бернштейн Н.А.* О построении движений. М.: Медгиз. 1947.

Bridgeman 2015 – Bridgeman B. (2015). *Consciousness*. NY: Oxford Bibliographies.

Burgess, Cohen-Yaacovi, Volle 2012 — Burgess P. W., Cohen-Yaacovi G., Volle E. (2012). Rostral prefrontal cortex. In B. Levine, F. I. M. Craik (eds) *Mind and the frontal lobes* (pp. 47–92). NY: Oxford University Press.

Casali et al 2013 – Casali A.G., Gosseries O., Rosanova M. et al. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine*, 5, 198ra105.

Chalmers 1996 – Chalmers D. J. (1996). *The conscious mind*. NY: Oxford University Press.

Crick, Koch 2005 – Crick F., Koch Ch. (2005). What is the function of the claustrum? *Phil. Trans. R. Soc. Lond. B. Biol. Sci.*, 360(1458), pp. 1271-1279.

Dehaene, Lau, Kouider 2017 — Dehaene, S., Lau, H., & Kouider, S. (2017). What is consciousness, and could machines have it? *Science*, 358, 486–492

Dennett 1992 – Dennett D. (1992). *Consciousness explained*. Boston: Little, Brown.

Dennett 2003 — Dennett D. (2003). *Freedom evolves*. NY: Viking Press, 2003

Descartes 1640 — Descartes R. (1640). Letter to Mersenne, 24 December 1640. In C. Adam, P. Tannery (eds.) *Oeuvres de Descartes* (vol. III, no. 223), 2nd ed., Paris: Vrin, 1964–1974.

Dickerson, Eichenbaum 2010 — Dickerson B. C., Eichenbaum H. (2010). The episodic memory system. *Neuropsychopharmacology Reviews*, 35, pp. 86-104.

Di Perri et al 2016 — Di Perri C., Ali Bahri M., Amico E., et al. (2016). Neural correlates of consciousness in patients who have emerged from a minimally conscious state. *Lancet Neurol.*, 15(8), pp. 830-842.

Dolina et al 2017 — Dolina, I. A., Efimova, O. I., Kildyushov, E. M., et al. (2017). Exploring *terra incognita* of cognitive science. *Psychology in Russia: State of the Art*, 10(3), pp. 231-247.

Edelman 1990 — Edelman G.M. (1990). *The remembered present: A biological theory of consciousness*. NY: Basic Books.

Fauconnier 2018 — Fauconnier G. (2018). *Ten lectures on cognitive construction of meaning*. Leiden/Boston: Brill

Fox, Christoff (eds) 2018 — Fox K.C.R., Christoff K. (eds) (2018). *The Oxford handbook of spontaneous thought*. NY: Oxford University Press.

Habib, Nyberg, Tulving 2003 — Habib R., Nyberg L., Tulving E. (2003). Hemispheric asymmetries of memory. *Trends in Cognitive Science*, 7(6), pp. 241–245.

Hameroff, Penrose 2014 — Hameroff S, Penrose R (2014). Consciousness in the universe. *Physics of Life Reviews*, 11, pp.:39–78.

He et al 2017 — He, Zh., Han, D., Efimova, O., et al. (2017). Comprehensive transcriptome analysis of neocortical layers in humans, chimpanzees and macaques. *Nature Neuroscience*. 20(6), pp. 886-895.

Huth et al. 2016 — Huth A. G., de Heer W. A., Griffiths T. L., et al. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532(7600), pp. 453-458.

Канеман 2017 — Канеман Д. *Думай медленно... Решай быстро!* М.: АСТ, 2017. Kahneman D. (2011). *Thinking, fast and slow*. NY: Farrar, Straus & Giroux.

Knyazev et al. 2017 — Knyazev G.G., Savostyanov A.N., Bocharov A.V. et al. (2017). Effortful control and resting state networks. *Neuroscience*, 346, pp. 365-381.

Kotov, 2017 — Kotov A. A. (2017). A computational model of consciousness for artificial emotional agents. *Psychology in Russia: State of the Art*, 10(3), pp. 57-73.

Lake et al. 2017 — Lake B. M., Ullman T. D., Tenenbaum J. B., Gershman S. J. (2017). Building machines that learn and think like people. *Behavioural and Brain Sciences*, 40, e253.

Lau, Rosenthal 2011 — Lau H., Rosenthal D. M. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences*, 15(8), pp. 365-374.

Лекторский 2010 — Лекторский В.А. «Я» // Новая философская энциклопедия (2-е изд.). М.: Мысль, 2010. Т. 4. С. 497-502.

Лекторский 2017 — Лекторский В.А. Являются ли иллюзией представления о субъективном мире? / Лекторский В. А. (ред.). *Субъективный мир в свете вызовов современной когнитивной науки*. М.: Аквилон, 2017. С. 7-21.

- Lektorsky 2017 — Lektorsky V. A. (2017). Realism as the methodological strategy in the cognitive science. In E. Agazzi (ed) *Varieties of scientific realism* (pp. 353–366). Berlin: Springer.
- Lisska 2016 — Lisska A.J. (2016). *Aquinas's theory of perception*. Oxford, UK: Oxford University Press.
- Liu, Perry (eds) 2014 — Liu J.L, Perry J., eds. (2014). *Consciousness and the Self*. Cambridge, MA.: Cambridge University Press.
- Лурия 1969 — Лурия А.Р. Высшие корковые функции человека и их нарушения при локальных поражениях мозга. М.: Издательство МГУ, 1969.
- Milner, Goodale 2008 — Milner A.D., Goodale M.A. (2008). Two visual systems re-viewed. *Neuropsychologia*, 46, pp. 774–785.
- Minsky 1988 — Minsky M. L. (1988), *The society of mind*. NY/ London: Touchstone Book.
- Moser E., Moser M. 2008 — Moser E. I., Moser M. B. (2008). A metric for space. *Hippocampus*, 18, pp. 1142-1156.
- Penfield, Evans 1935 — Penfield, W., & Evans, J. (1935). The frontal lobe in man: A clinical study of maximum removals. *Brain*, 58, pp. 115-133.
- Петренко, Супрун 2017 — Петренко В.Ф., Супрун А.П. Методологические пересечения психосемантики сознания и квантовой теории. М.: URSS.
- Петровский 2013 — Петровский В.А. “Я” в персонологической перспективе. М.: Издательство ВША, 2013.
- Raichle 2015 — Raichle M.E. (2015). The restless brain. *Phil. Trans. R. Soc. B* 370(1668): 20140172.
- Remedios, Logothetis, Kayser 2010 — Remedios R., Logothetis N. K., Kayser C. (2010). Unimodal responses prevail within the multisensory Claustrum. *Journal of Neuroscience*, 30(39): pp.12902–12907.
- Rosenthal 2005 — Rosenthal D. M. (2005). *Consciousness and mind*. Oxford: Clarendon Press.
- Rosenthal 2018 — Rosenthal D. M. (2018 in press). Consciousness and confidence. *Neuropsychologia*, doi: 10.1016/j.neuropsychologia.2018.01.018
- Sharaev et al. 2016 — Sharaev M.G., Zavyalova V.V., Ushakov V.L., et al. (2016). Effective connectivity within the default mode network. *Frontiers in Human Neuroscience*, 10, Art. No. 14.

Shishkin et al. 2016 — Shishkin, S.L., Nuzhdin, Y.O., Svirin E.P., & Velichkovsky, B.M. (2016). EEG Negativity in fixations used for gaze-based control: Toward converting intentions into actions with an Eye-Brain-Computer Interface. *Frontiers in Neuroscience*, 10, Art. No. 528.

Sloman, Chrisley, 2003 — Sloman A., Chrisley R. (2003). Virtual machines and consciousness. *Journal of Consciousness Studies*, 10(4–5), pp. 133-172.

Соколов 2010 — Соколов Е.Н. Очерки по психофизиологии сознания. М.: Издательство МГУ, 2010.

Striedter 2016 — Striedter G.F. (2016). Evolution of the hippocampus in reptiles and birds. *Journal of Comparative Neurology*, 524(3), pp. 496-517.

Tononi 2012 — Tononi G. (2012). The integrated information theory of consciousness: An updated account. *Archives Italiennes de Biologie*, 150, pp. 56–90.

Valitutti, Trautteur 2017 — Valitutti A., Trautteur G. (2017). Providing self-aware systems with reflexivity. *Lecture Notes in Computer Science*, 10640 LNCS, pp. 418–427.

Величковский 2015 — Величковский Б.М. Сознание // *Большая Российская Энциклопедия*. М.: Большая Российская Энциклопедия, 2015. Т. 30, С. 623-626.

Величковский 2017 — Величковский Б.М. От исследований сознания к разработке когнитивных технологий / Лекторский В. А. (отв. ред.). Субъективный мир в свете вызовов современной когнитивной науки. М.: Аквилон, 2017. С. 37-57.

Величковский, Ковальчук, Ушаков, Шараев 2016 — Величковский Б.М., Ковальчук М.В., Ушаков В.Л., Шараев М.Г. Исследование сознания естественнонаучными методами // *Вестник РФФИ*. 2016. № 91(3). С. 61-71.

Величковский и др. 2016 — Величковский Б.М., Нуждин Ю.О., Сvirin Е.П. и др. Управление «силой мысли»: на пути к новым формам взаимодействия человека с техническими устройствами // *Вопросы психологии*. 2016. № 62(1). С. 79 - 88.

Velichkovsky 2002 — Velichkovsky B.M. (2002). Heterarchy of cognition: The depth and the highs of a framework for memory research. *Memory*, 10 (5/6), pp. 405-419.

Velichkovsky 2005 — Velichkovsky B.M. (2005). Modularity of cognitive organization: Why it is so appealing and why it is wrong. In W. Callebaut, D. Rasskin-Gutman (eds.), *Modularity* (pp. 335-356). Cambridge, MA: MIT Press.

Velichkovsky et al. 2005 — Velichkovsky B.M., Joos M., Helmert J.R., Pannasch S. (2005). Two visual systems and their eye movements. In B. G. Bara, L. Barsalou, M. Bucciarelli (eds.), *Proceedings of the XXVII Conference of the Cognitive Science Society* (pp. 2283-2288). Mahwah, NJ: Erlbaum.

Velichkovsky et al. 2017 — Velichkovsky B.M., Krotkova O.A., Sharaev M.G., Ushakov V.L. (2017). In search of the “I”. *Psychology in Russia: State of the art*. 10(3), pp. 7-27.

Velichkovsky et al. 2018 — Velichkovsky B.M., Krotkova O.A., Kotov A.A., et al. (2018 in press). Consciousness in a multilevel architecture. *Consciousness and Cognition*. doi: 10.1016/j.concog.2018.06.004

Велихов, Зинченко, Лекторский 1988 — *Велихов Е.П., Зинченко В.П., Лекторский В.А.* Сознание: опыт междисциплинарного подхода // Вопросы философии. 1988, № 11. С. 3-30.

Vogeley 2017 — Vogeley K. (2017). Two social brains: Neural mechanisms of intersubjectivity. *Phil. Trans. R. Soc. B* 372: 20160245.

Зинченко 2010 — *Зинченко В.П.* Сознание и творческий акт. М.: Языки славянских культур, 2010.