

## **Байесовский разум: новая перспектива в когнитивной науке\***

**М.А. Сущин**

В статье обсуждается новая амбициозная программа в современных когнитивных исследованиях, характеризующая перцептивные механизмы разума как вероятностную статистическую машину, пытающуюся минимизировать расхождения между своими предсказаниями положения дел в среде и идущими из мира сигналами. Автор утверждает, что данная программа предоставляет наиболее систематический способ дополнить экологическую стратегию в исследованиях восприятия и познания, развитую в работах Дж. Дж. Гибсона и приверженцев современного энактивизма.

**КЛЮЧЕВЫЕ СЛОВА:** Байесовский мозг, внимание, восприятие, действие, первичные ожидания, предсказывающая обработка, прямая модель.

**СУЩИН** Михаил Александрович – кандидат философских наук, старший научный сотрудник кафедры философии и социологии Юго-Западного государственного университета.

### **The Bayesian Mind: A New Theory in Cognitive Investigations**

**M.A. Sushchin**

This article discusses a new intriguing possibility in contemporary cognitive investigations which proposes that the perceptual machinery of the mind is a probabilistic knowledge-driven inference machine that tries to reduce discrepancies between its predictions and actual sensory signals. Though lacking sufficient empirical evidence for its full justification, Bayesian modeling is still the most systematic way to approach old significant problems of cognitive science and, as I will argue, it is necessary to complement an ecological and actionist perspective elaborated by J. J. Gibson and modern enactivists.

**KEYWORDS:** action, attention, Bayesian brain, forward model, perception, predictive processing, prior expectations.

**SUSHCHIN** Mikhail A. – CS in Philosophy, Senior Research Fellow of Department of Philosophy and Sociology of Southwest State University.

sushchin@bk.ru

### **Введение**

---

\* Работа выполнена при поддержке гранта РФФ, проект № 15-18-10013 «Социо-антропологические измерения конвергентных технологий». This work was supported by the Russian Science Foundation, project No. 15-18-10013 “Socio-anthropological dimensions of converging technologies”.

Новая масштабная и перспективная теория стремительно обретает популярность в современных исследованиях познания и мозга. Известная преимущественно как «Предсказывающая обработка/Предсказывающее кодирование» (англ. «Predictive processing/Predictive coding») [Clark 2013a] или же «подход минимизации ошибки в предсказании» [Hohwy 2013, 49, 65, 147], данная теория, по утверждению одного из авторов, претендует на звание «наиболее полной к настоящему моменту концептуальной рамки для объяснения восприятия, познания и действия в терминах фундаментальных теоретических принципов и нейрокогнитивных архитектур» [Seth 2015 web]. Кроме того, в последнее время на основе ее ключевых принципов были предложены попытки объяснения (по крайней мере, частичного) столь разнообразных когнитивных, психофизических и ментальных феноменов, как внимание [Hohwy 2013, 191–206], иллюзии, психические расстройства [Fletcher, Frith 2009], опыт, сознание, Я [Hohwy 2013] и эмоции.

В центре данной амбициозной программы исследований находится простая, но вместе с тем потенциально объяснительно мощная идея, а именно что перцептивные механизмы разума представляют собой вероятностную нагруженную знанием статистическую машину, которая стремится минимизировать разницу между своими предсказаниями положения дел в среде и идущими из внешнего мира сигналами. Важной посылкой, разделяемой большинством исследователей в этой области, является то, что перцептивные процессы оперируют в условиях существенной неопределенности, имеющей своим источником неопределенность воспринимаемых стимулов и/или шум/случайные ошибки в процессе нейронной обработки сигналов, результаты которого, с этой точки зрения, и стремится нивелировать перцептивная система. Поэтому, восприятие, рассмотренное с позиции реализующих его публичностных процессов и механизмов, согласно сторонникам байесовской программы, является, в сущности, не чем иным, как вероятностным выводом.

Данная идея имеет долгую и богатую историю. Так, еще арабский ученый-энциклопедист XI века Ибн аль-Хайсам (Альхазен) придерживался точки зрения, что «многие видимые свойства воспринимаются посредством суждения и вывода» [Hohwy 2013, 5]. Кант, постулировавший априорные структуры (априорные формы пространства и времени и априорные категории рассудка) для оформления предоставляемого чувственностью материала [Кант 1994], также иногда рассматривается в качестве одной из предтеч данного направления. С другой стороны, основоположниками математической теории вероятностей Т. Байесом и П.С. Лапласом было введено строгое формальное правило для интерпретации новых данных с учетом уже имеющегося знания – так называемое правило Байеса (или теорема Байеса)<sup>1</sup>, активно используемое ныне в качестве модели работы процессов восприятия и познания.

И все же ключевая роль в развитии этого направления отводится знаменитому немецкому врачу, физику и физиологу Герману фон Гельмгольцу. Именно Гельмгольц [Helmholtz 1962] впервые в ясном виде сформулировал представление о восприятии как об (индуктивных) умозаклучениях, производимых нервной системой без участия сознания. «Психические

активности, – указывал он, – которые ведут нас к заключению, что перед нами в определенном месте существует определенный объект определенного характера являются в общем не сознательными активностями, а бессознательными. По их эффектам они аналогичны *умозаключениям* (курсив автора – М.С.) в той мере, что наблюдаемое действие на наши чувства позволяет нам формировать идею возможной причины этого действия; хотя на самом деле просто неизменно нервные стимуляции воспринимаются прямо, т.е. действия, но никогда не сами внешние объекты» [Ibid., 4].

Помимо этого, Гельмгольц по аналогии со значением экспериментов в науке в весьма современном ключе подчеркивал роль и необходимость действий и двигательной активности, которые, как он указывал, играют неопределимую роль в усилении корректности суждений о причинах наших восприятий. В XX веке заложенная Гельмгольцем конструктивистская традиция в исследованиях восприятия была продолжена в работах таких известных исследователей, как Р. Грегори [Gregory 1980; Gregory 1997], И. Рок [Rock 1983], У. Найссер [Neisser 2014], Дж. Брунер [Брунер 1977], Г. Каницца [Kanizsa 1985] и др.

Другая важная составляющая центрального постулата предсказывающего кодирования, идея минимизации (или коррекции) ошибки в предсказании (англ. prediction error), как указывает философ Энди Кларк [Clark 2013a, 182–183], была концептуально навеяна разработанными в 1950-х гг. инженерами Лаборатории Белла техниками сжатия данных в процессе передачи и обработки сигналов. Например, для передачи большого объема данных (предположим, большого изображения) было предложено использовать кодирование только тех его частей (тех самых «ошибок в предсказании»), которые отличаются от типичных и не ожидаются получателем – все остальное должно быть успешно предсказано и декодировано надлежащим образом информированным получателем, позволяя достичь существенной экономии ресурсов каналов связи (наследники данной стратегии, отмечает Кларк, в настоящее время используются в JPEG и других методах сжатия данных).

Коротко говоря, в начале 1990-х гг. исследователи из университетов США и Великобритании, работавшие на стыке математики и психологии, опираясь на байесовскую теорию принятия решений и методы математической статистики, начали разрабатывать детальные байесовские модели восприятия, перцептивных процессов и сенсорно-моторного контроля [Rescorla 2015 web]. Постепенно, аккумулируя на этом фундаменте идеи и принципы многих областей и дисциплин (включая кибернетику, теорию информации, генеративные подходы в машинном обучении и теории искусственных нейронных сетей), приверженцы возникшей программы, как уже говорилось в самом начале, стали в той или иной степени применять байесовские конструкты для объяснения широкого круга когнитивных и психофизических феноменов и процессов.

В настоящей статье мы планируем рассмотреть основные принципы байесовской программы в когнитивных исследованиях, преимущественные области ее приложения, перспективы развития, наиболее острые проблемы, а

также отношение к некоторым другим известным традициям в когнитивной науке. В частности, особый интерес представляет отношение возникшего поля исследований к традиции, восходящей к работам Дж. Дж. Гибсона, а в современной литературе представленной исследованиями сторонников так называемого «ситуативного/воплощенного познания» [Сущин 2014] и «энактивизма» [O'Regan, Noe 2001]. При всей революционности исследований Гибсона [Лекторский 2009, 31] его подход часто обоснованно критиковался за парадоксальное отсутствие внимания к роли «внутренних» психофизиологических факторов в восприятии и познании. Байесовская теория восприятия на текущий момент занимает наиболее видные позиции среди теорий, пытающихся отдать должное этим факторам.

### **Основные принципы: функциональная инверсия прямой и обратной сенсорной связи**

Настоящее рассмотрение мы начнем с наиважнейшей части байесовской программы в исследованиях познания, посвященной анализу проблемы восприятия (прежде всего, конечно же, внешнего мира, или экстероцепции). Итак, в основании данного направления исследований лежит совокупность теоретических принципов, радикальным образом отличающих его от другой известной традиции в изучении восприятия, получившей распространение в нейронауке, когнитивной психологии и исследованиях компьютерного зрения в 1960–1980-х гг. (к сторонникам этого лагеря относят, прежде всего, знаменитых нейрочеловеческих Нобелевских лауреатов Д. Хьюбеля и Т. Визеля, а также известных ученых в области исследований нейронауки и машинного зрения Д. Марра [Marr 1987] и И. Бидермана [Biederman 1987]). Данная более традиционная модель восприятия, взятая в несколько упрощенном виде, представляет его как в существенной степени восходящий (англ. bottom-up) последовательный процесс аккумуляции и объединения свойств и элементов зрительных сцен (базовых примитивов изображений, таких, как линии, границы, распределение яркости, группирующихся и организующихся в формы и отношения по мере распространения сигнала вверх по нейровычислительной иерархии от начальных сенсорных «станций» вроде сетчатки и таламуса на пути к высшим регионам коры мозга). Лишь «На определенном этапе эти сложные формы и отношения активизируют массивы хранящегося знания, превращая поток ощущений в раскрывающие мир (world-revealing) восприятия: видение кофе, пара и чашки, с паром, ассоциируемым с кофе, красным цветом, связываемым с чашкой и т.д. Назовем эту модель перцептивных процессов «пассивной аккумуляцией», – пишет известный философ и когнитивист Энди Кларк [Clark 2013b web].

Совершенно иначе представляет весь процесс предсказывающая нисходящая (англ. top-down) модель восприятия, основные принципы которой были изложены в классических работах знаменитого математика Дэвида Мамфорда [Mumford 1992; Lee, Mumford 2003], работе нейрочеловеческих Р. Рао и Д. Балларда [Rao, Ballard 1999] (а также работах некоторых других авторов [Kawato, Hayakawa, Inui 1993]). С этой (возможно, контринтуитивной) точки

зрения, многоуровневая иерархия перцептивных механизмов разума/мозга еще до столкновения с начальным потоком входящей сенсорной информации располагала определенными сенсорными ожиданиями (или, как их также часто называют, априорными ожиданиями – англ. *prior expectations*) относительно того, с чем может иметь дело непосредственно сенсорный вход.

Например, согласно модели взаимодействия высших и низших регионов коры мозга Д. Мамфорда, для связанных между собой регионов коры иерархически более высокоуровневый регион *A* будет генерировать предсказания (в терминологии Мамфорда, шаблоны) относительно того, с какого рода сенсорной информацией будет иметь дело более низкоуровневый регион *B* (например, лицо, дом, кот, дорожный знак и т.п.). Далее, высокоуровневый регион сообщает свое «видение» сенсорного входа нижестоящему слою, где оно переводится в «понятную» ему форму и сопоставляется непосредственно с поступившей от органов чувств текущей сенсорной информацией. И либо предсказание региона *A* с учетом определенного уровня шума является исчерпывающим, и вся система в целом почти совершенно предсказывает сенсорный вход, либо нижестоящий регион *B* будет вынужден вычислить «остаток, описание той части мира, которая не ожидалась или предсказывалась (т.е. ту самую пресловутую ошибку в предсказании – М.С.)» [Mumford 1992, 247]. Этот «остаток» (или ошибка в предсказании) затем отправляется наверх для внесения соответствующих корректив в гибкие шаблоны высокоуровневого региона с целью порождения более точных предсказаний воспринимаемых объектов. В случае же, если мы удивлены чем-либо и не имеем соответствующих предшествующих перцептивных ожиданий и гипотез, указывает Мамфорд, «алгоритм стартует с чистого листа в зоне *A*. Затем *B* просто отправляет *A* всю свою картину мира, которая возбуждает некоторые возможные высокоуровневые объекты. На каждом этапе *A* выписывает на своей доске (*blackboard*) свои лучшие догадки на своем собственном языке (объекты и их параметры) о характере высокоуровневых объектов, найденных в картине *B*» [Ibid.].

Поразительной особенностью предложенной Мамфордом когнитивной архитектуры, сообщившей импульс для последующих аналогичных теоретических разработок, является фактическая инверсия функциональных ролей систем прямой и обратной сенсорной связи в мозге. Это означает, как свидетельствуют развившие эту тему адепты предсказывающей обработки, что восприятие в содержательном плане почти всецело определяется потоком нисходящих предсказывающих сигналов, пошагово спускающихся по «ступеням» перцептивной нейровычислительной иерархии (где каждый вышестоящий уровень пытается предсказать активность ближайшего ему нижнего уровня) и блокирующих любую восходящую информационную активность, кроме той, которая содержит неучтенные нисходящими предсказаниями аспекты среды (т.е. те самые ошибки в предсказании). Стало быть, именно иницируемая разумом/мозгом нисходящая вычислительная активность, если предположить, что данная модель является корректной, в функциональном отношении предстает как прямая, в то время как идущая из

мира активность выступает здесь как источник обратной связи на предсказания генеративной модели мозга.

Как уже отмечалось нами в самом начале статьи, согласно ключевой исходной посылке современных байесовских моделей, перцептивные процессы типичным образом происходят в условиях существенной неопределенности, будь то, как утверждается, неопределенность, связанная с шумом/случайными ошибками в процессе сенсорного кодирования сигналов, или же неопределенность непосредственно воспринимаемых сцен и объектов (возникающая, например, когда различные модальности могут предоставлять нуждающуюся в интеграции противоречивую информацию). В полном согласии с идеями Гельмгольца приверженцы современных байесовских подходов исходят из предположения о непрямом характере перцептивных процессов: перцептивная система не имеет прямого доступа к воздействовавшим на ее вход «скрытым причинам» (англ. «hidden causes»). Все, чем располагает мозг и нервная система, есть лишь поступившие на вход сигналы, которые подлежат сопоставлению с производимым генеративной моделью мозга потоком нисходящих предсказывающих сигналов. Чтобы порождаемые разумом/мозгом предсказания в лучшей степени соответствовали входящему потоку сигналов, мозг, согласно общепринятой среди сторонников предсказывающей обработки модели, вынужден полагаться на процессы иерархического байесовского вывода, осуществляющие отбор наилучшей гипотезы для предсказания текущего сенсорного входа.

Далее, предполагается, что весь процесс селекции гипотез происходит в примерном соответствии с упомянутым ранее правилом Байеса<sup>2</sup>, которое представляет собой формальный метод вычисления условных вероятностей (т.е. вероятности какого-либо события  $A$  при условии наступления события  $B$ ; например, вероятности того или иного заболевания при условии наличия определенных симптомов). Обобщенно говоря, смысл правила Байеса (или теоремы Байеса) заключается в определении апостериорной вероятности гипотезы с учетом имеющихся у нас данных (функция правдоподобия) и, прежде всего, нашего предшествующего знания о том, насколько данная гипотеза вероятна в принципе. Мы используем наше предшествующее знание (в формальном отношении известное как априорная вероятность) для интерпретации новых данных и итоговом приписывании апостериорной вероятности интересующей нас гипотезе.

Например, нас может интересовать вопрос о том, является ли присутствующий на зрительной сцене объект выпуклым или вогнутым. В таком случае определенный зрительный вход  $I$  (данные) может быть совместим как с гипотезой выпуклого, так и с гипотезой вогнутого объекта. Тогда в дело вступает априорная вероятность, которая говорит, что *ceteris paribus* мозг склонен исходить из предположения, что источник света должен располагаться над головой и что в этом случае существует более высокая степень вероятности, что находящийся перед ним объект является выпуклым (что, соответственно, влечет за собой приписывание данной гипотезе более высокой апостериорной вероятности). Наконец, гипотезе, получившей наибольшую

апостериорную вероятность (так называемый апостериорный максимум), согласно байесовской модели, допускается контролировать нисходящую динамику процесса восприятия и предсказывать эволюцию сенсорного входа при дальнейшем взаимодействии с объектами.

Так, например, если была принята гипотеза, что на зрительной сцене в данный момент находится дом, то при смене перспектив мозг будет исходить из предположения, что данный объект на самом деле является домом, а не, скажем, только фасадной стеной, за которой ничего не скрывается. Принятые гипотезы являются чувствительными к текущей обстановке и доходящим из мира сигналам в интерактивной манере – если во время осмотра зрительной сцены «гипотеза дома» постепенно получает подтверждение, то ее вероятность возрастает все более, и наоборот, вызванный гипотезой поток ошибок инициирует новый этап селекции гипотез и т.д.

Нужно сказать, что представленная выше схема является, конечно же, в высокой степени идеализированной. Так, начальное пространство гипотез, из которого мозг, согласно сторонникам этой модели, должен выбирать гипотезы для предсказания сенсорного входа, для множества ситуаций является неисчислимым [Rescorla 2015 web], что делает байесовскую схему неправдоподобной в качестве действительного алгоритма, которому мог бы следовать мозг. В конце концов, отнюдь не ясен общенаучный статус утверждений, что мозг или разум могут осуществлять вычисления того или иного рода. Согласно одной точке зрения, «Байесовская модель только приблизительно описывает *действительные* (здесь и далее курсив автора – М.С.) ментальные процессы. Модель пытается описать, возможно, идеализированным образом, как перцептивная система *в действительности* переходит от сенсорного входа к перцептивным оценкам», – отмечает специалист в области философии математики Майкл Рескорла [Ibid.].

Впрочем, подобного рода схемы все же широко привлекались для объяснения целого круга перцептивных, сенсорно-моторных и когнитивных феноменов, включая мультисенсорную интеграцию (интеграцию информации, поступающей от различных сенсорных модальностей), бистабильное восприятие (куб Неккера, бинокулярное соперничество и т.п.), а также действие и моторный контроль, полное освещение и оценка которых выходит далеко за рамки настоящего рассмотрения.

### **Действие: активный вывод vs. оптимальный контроль**

Приведенный выше панорамный обзор основных принципов предсказывающей обработки был сфокусирован лишь на одной стороне общего механизма минимизации ошибки в предсказании, а именно восприятии. В соответствии с представленной картиной, разум и мозг руководствуются (или действуют в примерном соответствии с) процессами иерархического байесовского вывода для выбора гипотез с целью предсказания входящих сенсорных сигналов и, собственно, порождения перцептивного опыта как такового. Если же, как мы видели, предсказания системы оказываются неточны, входящий поток ошибочных сигналов используется, чтобы внести

необходимые коррективы во внутреннюю генеративную модель мира для установления лучшего соответствия между ее гипотезами и самим миром.

Очевидно, что данное видение процессов минимизации ошибок в предсказании нуждается в надлежащей оценке и учете роли действия и двигательной активности, поскольку тривиальное обстоятельство заключается в том, что мы в максимально возможной степени используем потенциал своих двигательных навыков, чтобы делать перцептивный мир в лучшей степени доступным нашим органам чувств и «... что мы, разумеется, используем способ нашего восприятия мира, чтобы информировать и направлять нашу способность действовать (agency)» [Hohwy 2013, 75–76].

Мы уже говорили, что еще основоположник всей конструктивистской программы в исследованиях восприятия Герман фон Гельмгольц подчеркивал значимость двигательной активности для усиления корректности производимых нервной системой бессознательных умозаключений о причинах воздействия на органы чувств. В этом аспекте своей теории Гельмгольц активно прибегал к аналогии со значением экспериментов в науке, в рамках которых, утверждал он, благодаря нашему непосредственному вмешательству и контролю условий возникновения интересующих нас событий, представляется единственно возможным установить надежную причинную связь между событиями. В противном случае, указывал Гельмгольц, если бы мы были лишены нашей способности активно исследовать мир при помощи действий и могли лишь пассивно созерцать, как объекты проходят перед нашим взором без малейшей возможности повлиять на них «... вероятно, мы никогда бы не смогли найти свой путь среди подобной оптической фантазмагии» [Helmholtz 1962, 31]. (Позднее к аналогичной метафоре, уподоблявшей восприятия проверяемым в науке гипотезам, обращался также и Ричард Грегори [Gregory 1978, 225; Gregory 1997, 1121].)

Таким образом, в рамках семейства байесовских подходов к когнитивным процессам в широком смысле были разработаны две основные теории действия, двигательной активности и связи сенсорных и моторных механизмов. Однако, прежде чем переходить к их рассмотрению, необходимо оговориться, что в рамках этих теорий действие и моторные процессы исследуются под весьма отличным углом, нежели в другой известной когнитивистской традиции, восходящей к работам М. Мерло-Понти и Дж. Дж. Гибсона и достигшей кульминации в исследованиях приверженцев того самого «ситуативного и воплощенного познания» [Сущин 2014, 44–57] и «энактивизма» [Noe 2004; Сущин 2015, 191–210].

С точки зрения последних, речь могла идти о фактическом размывании границ между восприятием, действием и познанием в целом или же о том, что, например, перцептивное сознание может определяться с качественной стороны владением особым сенсорно-моторным знанием (сенсорно-моторными зависимостями), благодаря которым, как утверждают адепты энактивизма, мы и в состоянии воспринимать мир [O'Regan, Noe 2001]. С другой стороны, в рамках характерной для байесовских теорий проблематизации действия и моторных процессов речь идет скорее о том, что Дэвид Марр в рамках своей



методологической схемы анализа систем обработки информации называл общей целью вычислительного процесса [Марр 1987]: байесовские модели принципиальным образом стремятся прояснить общую цель моторной активности (например, ее роль в рамках механизма минимизации ошибки в предсказании), фактически оставляя вопрос о конкретных алгоритмах и деталях физической реализации вне рамок своей компетенции.

Итак, первая интересующая нас байесовская модель моторных процессов была разработана как часть программы «предсказывающей обработки» и получила (по аналогии с базовыми процессами перцептивного вывода) наименование «активного вывода» (англ. «active inference»). Ее смысл, обобщенно говоря, заключается в том, что когнитивная система, как постулируется, способна минимизировать ошибки в предсказании не только посредством пересмотра и обновления внутренней генеративной модели мира (когда ее предсказания оказываются неверны), но и при помощи двигательной активности и действий (например, локомоций), а именно через изменение своего положения в мире и/или активное воздействие на текущее положение вещей, чтобы сенсорный вход в лучшей степени соответствовал предсказаниям модели [Hohwy 2013, 77]. (Например, чтобы устранить двусмысленность наблюдаемой сцены и/или увеличить достоверность определенной перцептивной гипотезы.)

С этой точки зрения, восприятие и действие оказываются двумя гранями единого механизма минимизации ошибок в предсказании сенсорного входа, подчиненными общей важнейшей стратегии и координирующимися для ее реализации в примерном соответствии со следующей схемой. Вначале перцептивная система просто-напросто следует описанным выше механизмам отбора гипотез для предсказания сенсорного входа, останавливается на гипотезе, получившей наибольшую апостериорную вероятность, и пытается предсказать, как сенсорный вход может изменяться, исходя из того, что эта гипотеза является верной. (Допустим, если была принята гипотеза дома, система пытается предсказать, что она сможет увидеть благодаря двигательной активности с недоступных с данной позиции перспектив.) Поскольку в данный конкретный момент положение вещей отличается от предсказанного (т.е. от моделей возможных зрительных сцен с учетом принятой гипотезы), возникает та самая пресловутая ошибка в предсказании, единственный способ устранения которой заключается в попытке реализации предсказаний модели с целью их подтверждения или опровержения. Если, как уже говорилось, гипотеза находит подтверждение, ее вероятность усиливается. В противном случае система возвращается к механизмам «пассивного» перцептивного вывода и пытается выбрать иную более подходящую гипотезу. Таким образом, с этой точки зрения, «действие не осуществляется через определенное сложное *вычисление* (курсив мой – М.С.) моторных команд, которые контролируют мускулы тела. Попросту говоря, напротив, происходит то, что мускулы двигаются до тех пор, пока существует ошибка в предсказании. Мускулы тела, поэтому, находятся во власти ошибки в предсказании, возникающей вследствие того, что мир не является таким, каким он должен быть в соответствии с моделью мозга.

Ошибка в предсказании в таком случае является простым механизмом, который управляет действием» [Ibid., 81].

В отличие от подобного рода механистического видения недавно возникшей концепции «активного вывода», более ранняя ветвь байесовского моделирования двигательной активности, известная как «оптимальный моторный контроль», рассматривала задачу реализации действия именно как вычислительную активность, основанную на использовании нескольких типов внутренних моделей. В этом отношении программа оптимального контроля «расширяет концепцию Гельмгольца до моторного контроля, постулируя бессознательные выводы, которые оценивают состояния среды, и процессы бессознательного принятия решений (unconscious decisions), которые выбирают моторные команды» [Rescorla 2016 web].

Ключевым здесь является изначально развитое в области управления сложными системами понятие прямой модели (англ. forward model). В общем виде предполагается, что контроль и управление действиями (с вычислительной точки зрения) происходит в соответствии со следующей схемой: вначале система на сознательном уровне устанавливает цель действия (допустим, взять стоящую на столе чашку чая). Затем (уже, разумеется, на субличностном уровне, при помощи так называемой обратной модели) вычисляется необходимая траектория движений для достижения установленной цели, которая посылается эффекторам (непосредственным физиологическим исполнителям действия). При этом на пути к эффекторам копия моторных команд (эфферентная копия), как предполагается, направляется той самой прямой модели, которая еще до появления действительных сигналов обратной связи на моторные команды системы создает своего рода вычислительную симуляцию, пытающуюся предвосхитить содержание сигналов будущей обратной связи на эту команду. Смысл подобной процедуры заключается в том, что система нуждается в максимально надежной оценке результатов своих действий, в то время как сигналы обратной связи часто бывают либо запоздалыми, либо излишне подверженными шуму/искажениям. Таким образом, «Главное преимущество прямых моделей заключается в том, – замечает признанный классик исследований моторного познания нейрочеловек Марк Жанро, – что они способны оценить желаемое состояние машины еще до ее действия» [Jeannerod 2006, 20].

Каким образом с этой точки зрения система останавливается на какой-либо одной траектории для реализации моторной цели из бесконечного множества всех возможных траекторий и почему не бывает двух абсолютно идентичных движений даже для одного и того же агента? Коротко говоря, моторная система, как постулируют сторонники данной модели, выбирает наиболее оптимальную траекторию, исходя из своей оценки положения дел в среде и отслеживает отклонения от предписанной последовательности команд, которые могут произойти от того же «шума, усталости или внешнего вмешательства. Всякий раз, когда случается отклонение, система сталкивается с выбором: скорректировать отклонение или игнорировать его.... Оптимальная стратегия заключается в том, чтобы скорректировать отклонение от средней

траектории, только если отклонение релевантно цели задачи» [Rescorla 2016 web].

Именно способность теории оптимального моторного контроля объяснить (или, по крайней мере, предложить определенное объяснение), как моторная система выбирает единственную уникальную траекторию для реализации цели, была недавно использована сторонниками этого направления как аргумент в споре с приверженцами альтернативной модели «активного вывода», в рамках которой, как утверждается, проблема избыточности степеней свободы [Бернштейн 1997] вовсе не получает адекватного рассмотрения [Rescorla 2016 web]. Так или иначе, но в рамках настоящего рассмотрения нет ни малейшей возможности рассудить спор этих двух направлений. Наша цель состояла всего лишь в том, чтобы продемонстрировать потенциал общей байесовской теории в свете непреложного факта тесной взаимозависимости систем восприятия и действия.

### **Внимание и другие области приложения**

Мы уже неоднократно подчеркивали, что восприятие и действие, безусловно, представляют собой центральный объект интереса для байесовских моделей в когнитивной науке, однако на попытках объяснения сенсорно-моторного конгломерата их амбиции отнюдь не ограничиваются. Так, недавно получили развитие попытки с позиции байесовского подхода раскрыть возможную роль внимания и специфики/источников некоторых психических расстройств (прежде всего, шизофрении).

В основании байесовского подхода к объяснению функций внимания лежит следующая мотивация: система минимизации ошибки в предсказании сенсорных сигналов функционирует в мире, изобилующем шумом или неопределенными сигналами. Поскольку пресловутые ошибочные сигналы играют неопределимую роль в обучении мозга и совершенствовании порождаемых его моделью предсказаний сенсорного входа, очевидно, что система должна располагать надежным механизмом фильтрации и отсеивания не имеющего значения шума от действительных ошибочных сигналов. Именно с выполнением этой важной задачи было предложено связать процессы внимания, которые при обнаружении подлинных ошибочных сигналов ранжируют их в соответствии со степенью точности, наделяя более точные сигналы большим весом и позволяя им таким образом распространяться до верхних «этажей» перцептивной иерархии и вносить необходимые коррективы в модель мира, чтобы она могла генерировать более точные и адекватные репрезентации положения дел в среде.

В связи с данной оригинальной интерпретацией функций внимания известным нейрочеловеком и популяризатором науки Крисом Фритом и его соавтором Полом Флетчером была предложена интересное объяснение происхождения и природы такого серьезного психического расстройства, как шизофрения [Fletcher, Frith 2009]. Так, замечают Фрит и Флетчер, «Наиболее разительными и характерными особенностями расстройства являются галлюцинации и ложные выводы (delusions). Галлюцинации являются ложными

восприятиями такого рода, что пациенты слышат других людей, разговаривающих о них, или слышат их мысли, [как будто] проговаривающиеся вслух. Ложные выводы есть устойчивые причудливые или иррациональные убеждения, которые не поддаются простому пониманию в терминах социального или культурного фона индивида. Например, пациенты могут верить, что другие люди способны слышать их мысли или что правительство следит за каждым их действием» [Ibid., 48]. Данные «позитивные симптомы» также усугубляются чувством «пассивности», когда индивидам кажется, что их действиями на самом деле руководят или манипулируют другие, а также чрезмерным вниманием к совершенно несущественным окружающим их предметам или событиям.

Фрит и Флетчер полагают, что источники подобной симптоматики могут быть рассмотрены сквозь призму понятийного аппарата байесовской теории восприятия и программы минимизации ошибки в предсказании. Поскольку, как уже неоднократно отмечалось, с этой точки зрения, не только не существует непроходимой границы между низкоуровневыми сенсорными механизмами и высокоуровневыми структурами знания и интеллектом, но и наоборот, как утверждается, структуры знания в существенной степени определяют сами восприятия и их содержание [Gregory 1997], то и характерные для шизофрении галлюцинации (восприятие) и ложные выводы (знание), полагают Фрит и Флетчер, могут быть связаны с нарушениями в работе общего механизма коррекции ошибок в предсказании.

Ключевая идея здесь заключается в следующем. Из исследований нейробиологических оснований шизофрении известно, что данное расстройство может быть обусловлено нарушениями работы системы, регулирующей активность нейромодулятора дофамин (который предположительно отвечает за оценку точности сигналов об ошибке, т.е. за работу процессов внимания, согласно представленной выше гипотезе). Рассуждая в терминах байесовской модели восприятия, это означает, что в работе механизма оценки точности сигналов об ошибке происходят нарушения, когда незначительным или ложным ошибочным сигналам присваивается несоразмерно большой вес, что коррелирует с упомянутой выше склонностью больных шизофренией к излишнему сосредоточению на совершенно несущественных вещах (например, на приоткрытом окне или замочной скважине). Далее этому ложному сигналу об ошибке в силу его неверно оцененной значимости дозволяется распространяться на верхние уровни перцептивной иерархии и вносить «роковые корректировки» в модель мира агентов. Все это запускает самоподдерживающийся порочный цикл, где искаженная модель мира будет генерировать неадекватные предсказания и интерпретации положения дел в среде, приводя, таким образом, к образованию столь характерных для шизофрении видов опыта.

Флетчер и Фрит полагают, что язык байесовской теории необходим для наведения концептуального «моста» между физическим (нарушениями работы подсистем мозга) и феноменальными (искаженным опытом в виде галлюцинаций и навязчивых бредовых состояний) уровнями в деле изучения

столь нуждающихся в объяснении феноменов человеческого опыта (тогда как любые изолированные исследования, подчеркивают авторы, неизбежно будут упираться в вопросы вроде: «Как дофамин может вызвать голос или убеждение?» [Fletcher, Frith 2009, 49]) Вопрос о возможном вкладе байесовской теории в объяснение феноменальных аспектов когнитивных процессов является, без сомнения, чрезвычайно важным, но, к сожалению, выходящим за рамки настоящего обсуждения. Сейчас же мы, прежде чем переходить к заключению, вкратце остановимся на еще одном важном аспекте байесовских моделей восприятия, а именно на предполагаемом ими характере отношения перцептивных процессов к миру.

### **Отношение к реализму и теориям прямого восприятия**

У заинтересованного читателя после всего сказанного выше мог возникнуть вопрос следующего характера: если байесовские модели восприятия постулируют, что содержание перцептивного опыта определяется гипотезой, получившей апостериорный максимум и успешно предвосхищающей поступающие из мира сигналы, в то время как сам мир всегда является скрытым от нас за завесой сенсорного входа, то не означает ли это, что байесовский подход в эпистемологическом плане оказывается разновидностью картезианского скептицизма или антиреализма, свойственного некоторым ранним моделям познания в когнитивной науке [Fodor 1981] и уступившего свои позиции с развитием исследований, подчеркивающих роль тела, воплощения, среды и социокультурных факторов в познании?

Именно к такому скептическому выводу в своих недавних работах склоняется один из ярких приверженцев байесовской программы в когнитивных исследованиях философ и когнитивист Якоб Хохви. Восприятие, утверждает он в полном согласии с идеями Гельмгольца, является непрямым и удаленным от мира на расстояние одного шага – сенсорного входа, за границами которого находятся скрытые причины, воздействующие на наши органы чувств и устанавливаемые мозгом в процессе его каузальных бессознательных выводов. Более того, настаивает Хохви, только с учетом такого понимания отношения перцептивных процессов к миру возможна сама программа минимизации ошибки в предсказании: «Это, – пишет он, – становится подтверждением простого Картезианского скептицизма. Поскольку мы не можем получить независимую от нашей позиции в мире точку зрения, мы не можем исключить скептическую гипотезу, что получаемый нами сенсорный вход вызван злонамеренным, обманывающим ученым, а не внешними положениями вещей, в которые мы обычно верим. Байесовская программа, поэтому, подразумевает скептицизм. Следовательно, отрицание скептицизма влечет за собой отрицание байесовского вывода, а значит, и ММО [механизма минимизации ошибки – М.С.]» [Hohwy 2014 web].

И хотя возможно, что в метафизическом отношении данная точка зрения является неопровержимой, мы исходим из того, что конкретная научная программа исследований не должна находиться в противоречии со связанными с ней метафизическими предпосылками. Более продуктивной и перспективной

в научном отношении нам представляется вовсе не гипотеза скептицизма или «методологического солипсизма» в духе Фодора, а неплохо зарекомендовавшая себя за последние четверть века в когнитивных исследованиях обычная реалистическая установка, постулирующая существование внешнего мира, вносящего ограничения на наши внутренние модели и знания. Ведь и сам Хохви в другом месте признает, что мир является истиной (хоть и доступной, по его мнению, только через завесу сенсорного входа), посредством посылаемых им ошибочных сигналов выступающей в роли естественного внешнего надзирателя (supervisor) процессов восприятия [Hohwy 2013, 48–51].

Мы, таким образом, полагаем, что байесовская программа в исследованиях восприятия и познания должна быть в фундаментальном смысле совместима с реалистической установкой и реалистически ориентированными исследованиями в когнитивной науке – прежде всего, конечно, тем же самым «ситуативным/воплощенным познанием». Более того, только реалистическая установка позволяет проложить дорогу к объяснению того, как формируются сами вычислительные механизмы разума и откуда берутся его априорные ожидания и гипотезы (что рассматривается в числе основных затруднений байесовских моделей восприятия). В противном случае, если кто-либо всерьез и вознамерился бы принять скептическую установку за ориентир для реализации своей научной программы, любые исследования эволюционных, телесных и социокультурных оснований познания (включая рассмотренную нами выше байесовскую теорию действия) для него просто потеряли бы актуальность – манипулируемый злонамеренным ученым «мозг в бочке», очевидно, не нуждается в двигательной активности, для него не существует ни социума, ни творимых им интеллектуальных артефактов и орудий.

Именно реалистическое понимание байесовских моделей когнитивных процессов было предложено в работах таких известных ее новоиспеченных сторонников, как Энди Кларк и Дэниел Деннет. Кларк, размышляя над отношением байесовской программы к «ситуативному/воплощенному познанию», высказал идею, что создание и эксплуатирование нами специальных интеллектуальных культурных сред в виде слов, дорожных знаков, книг, компьютеров и т.д., делая мир более удобным для нашего мозга, позволяет нам более эффективно минимизировать ошибки в предсказании в самых различных контекстах от повседневной жизни до специализированных видов деятельности вроде науки [Clark 2013a, 194–195].

Деннет предположил [Dennett 2013], что преимущественный набор (перцептивных) ожиданий организма, играющих ключевую роль в байесовской схеме, может быть связан с тем, что имеет для него наибольшую значимость и составляет его мир, или Umwelt, складывающийся в свою очередь из того, что Дж. Дж. Гибсон называл «возможностями» – коротко говоря, объектами среды, которые могут предоставить укрытие, быть употреблены в пищу и т.д. (т.е. в наиболее общем смысле принести пользу или вред).

Здесь может возникнуть возражение: но ведь Гибсон, долгое время отстаивавший точку зрения прямого (или наивного) реализма, считается едва ли не главным антиподом гельмгольцевского вычислительного понимания

восприятия – разве могут эти позиции быть хоть в малейшей степени объединены? (Особенно с учетом того обстоятельства, что Гибсон, как известно, подчеркивал, что восприятие возможностей является прямым.) На это мы должны еще раз заметить, что, хотя байесовская теория и расходится кардинально с теорией Гибсона по вопросу о характере отношения восприятия к миру, она все же является совместимой с реализмом в широком смысле. Иначе говоря, байесовская (иная вычислительная) программа выиграет больше, если будет отталкиваться от реалистической установки, а экологическая (ситуативная, воплощенная) программа – если не будет отрицать роль памяти и структур знания в процессах восприятия мира. Попытки примирить эти два лагеря предпринимались еще У. Найссером [Найссер 1981] и Д. Марром [Marr 1987].

### **Заключение: главные вызовы для байесовской программы в исследованиях познания**

Итак, в рамках настоящего обзора мы постарались рассмотреть ключевые принципы современных байесовских моделей восприятия, действия и познания. Байесовское моделирование когнитивных процессов является интригующим подходом к наиболее проблематичным проблемам когнитивной науки, уже успевшим получить поддержку многих известных ученых и философов, возлагающих большие надежды на его успешное развитие и разработку. Тем не менее, все это, разумеется, не означает, что возникшее направление вовсе свободно от каких бы то ни было затруднений или проблем – вовсе не так. Даже более того, рядом авторов была поставлена под сомнение принципиальная верность и перспективность курса, предлагаемого байесовскими моделями когнитивных процессов.

Так, в обстоятельном обзоре с говорящим названием «Байесовские сказки в психологии и нейронауке» психологи Дж. Бауэрс и К. Дэвис [Bowers, Davis 2012] детально рассмотрели и подвергли острой критике буквально каждое из ключевых положений байесовской программы в когнитивных исследованиях. С их точки зрения, байесовский подход к изучению разума и мозга является ошибочным, поскольку, во-первых, байесовские модели восприятия, моторного контроля и высокоуровневого познания, утверждают они, являются настолько гибкими и часто конструируемыми *post hoc*, чтобы «объяснить» фактически любые наблюдаемые данные, что, как следствие, на практике влечет их нефальсифицируемость. Например, в известных байесовских моделях восприятия скорости, распознавания слов и принятия решений ключевые компоненты теоремы Байеса (функция правдоподобия, априорная вероятность, а также добавляемая к ним так называемая функция полезности), как было убедительно показано в обзоре Бауэрса и Дэвиса, могут быть использованы произвольным образом, чтобы подстроиться практически под любые данные так, что «если бы оказалось, что данные являются другими, была бы построена отличная Байесовская модель, чтобы оправдать заключение, что работа человеческого познания (*performance*) является оптимальной» [Ibid., 410].

Во-вторых, что очень важно, как было верно замечено авторами, конкретные нейробиологические свидетельства в пользу того, что мозг в действительности по преимуществу следует байесовским вычислительным схемам, являются крайне слабыми (можно сказать, почти отсутствующими), а само положение об избыточном шуме в процессе сенсорного кодирования и нейронной обработки сигналов с учетом определенных экспериментальных данных – как минимум, нуждающимся в дополнительном подтверждении.

В-третьих, Бауэрсом и Дэвисом была оспорена лежащая в основании байесовского моделирования методология «рационального анализа», являющаяся, по сути, развитием программы исследований зрения Д. Марра, подчеркивавшего, как известно, приоритет и большую значимость уровня вычислительной задачи (т.е. общего смысла деятельности информационной или когнитивной системы) перед уровнями алгоритмической и физической реализации этой задачи. В ответ на это авторы резонно указали, что важнейшие для исследующих работу разума и мозга дисциплин свидетельства и ограничения могут быть продиктованы соображениями чисто эмпирического характера (например, анатомическим устройством сетчатки и зрительных путей в мозге, далекими от совершенства «эволюционными решениями» в общем), не мыслимыми и не выводимыми исключительно из рационального анализа задачи, которую выполняет система.

Резюмируя, нужно сказать, что авторам, безусловно, удалось наглядно показать слабые места байесовской программы в изучении восприятия и когнитивных процессов (особенно в аспекте чрезмерной общности и нефальсифицируемости байесовских моделей). В результате остаются открытыми вопросы о том, в какой степени работа когнитивных механизмов человека и других агентов в действительности подвержена шуму/искажениям и сопряжена ли она со следованием/реализацией вычислительных схем, постулируемых байесовскими моделями. Представляется, что решающий вклад в разрешение этих вопросов способны внести будущие нейрокогнитивистские исследования оснований работы сенсорно-моторных механизмов познания.

Однако пока эмпирические исследования не склонили чашу весов в определенную сторону, байесовская программа остается наиболее целостной и систематичной среди теорий, которые предлагают объяснение того, что происходит «в голове» когнитивных агентов (пусть в плане спецификации общей вычислительной задачи), а не только того, внутри чего эта голова находится [Mace 1977 web], т.е. воспринимаемого внешнего мира (как предполагалось Гибсоном и предполагается ныне во многом перенявшими гибсонизанскую установку приверженцами «энактивизма»). Или же в противном случае должны быть предоставлены всеобъемлющие доказательства, что то, что происходит «внутри черепной коробки», не имеет значения для понимания восприятия и познания. Насколько мы можем судить, до сих пор таких доказательств предоставлено не было [Суцин 2015].



Бернштейн 1997 – *Бернштейн Н.А.* Биомеханика и физиология движений / Под редакцией В.П. Зинченко. М.: Издательство «Институт практической психологии», Воронеж: НПО «МОДЭК», 1997.

Брунер 1977 – *Брунер Дж.* О готовности к восприятию // Брунер Дж. Психология познания. М.: Издательство «Прогресс», 1977. С. 13–64.

Кант 1994 – *Кант И.* Критика чистого разума. М.: Мысль, 1994.

Лекторский 2009 – *Лекторский В.А.* Реализм, антиреализм, конструктивизм и конструктивный реализм в философии и науке // Конструктивистский подход в эпистемологии и науках о человеке / Отв. ред. акад. РАН В.А. Лекторский. – М.: «Канон+» РООИ «Реабилитация», 2009. С. 5–40.

Марр 1987 – *Марр Д.* Зрение. Информационный подход к изучению представления и обработки зрительных образов. М.: Радио и связь, 1987.

Найссер 1981 – *Найссер У.* Познание и реальность: смысл и принципы когнитивной психологии. М.: Прогресс, 1981.

Сущин 2014 – *Сущин М.А.* Концепция ситуативного познания в когнитивной науке: критический анализ: дис. ... канд. филос. наук: 09.00.01 / Сущин Михаил Александрович: М., 2014.

Сущин 2015 – *Сущин М.А.* Вы – это Ваш мозг? Почему когнитивная нейронаука нуждается в поведенческой и ситуативной перспективе // Аршинов В.И. Асеева И.А., Буданов В.Г., Гребенщикова Е.Г., Гримов О.А., Каменский Е.Г., Москалев И.Е., Пирожкова С.В., Сущин М.А., Чеклецов В.В. Социо-антропологические измерения конвергентных технологий. Методологические аспекты: Коллективная монография / Отв. ред. И.А. Асеева, В.Г. Буданов. Курск: ЗАО «Университетская книга», 2015. С. 191–210.

## References

Biederman 1987 – *Biederman I.* Recognition-by-Components: A Theory of Human Image Understanding // *Psychological Review*. 1987. Vol. 92. No. 2. P. 115–147.

Bowers, Davis 2012 – *Bowers J. S., Davis C. J.* Bayesian Just-So Stories in Psychology and Neuroscience // *Psychological Bulletin*. 2012. Vol. 138. No. 3. P. 389–414.

Clark 2013a – *Clark A.* Whatever next? Predictive brains, situated agents, and the future of cognitive science // *Behavioral and Brain Sciences*. 2013. Vol. 36. No. 3. P. 181–204.

Clark 2013b web – *Clark A.* Expecting the World: Perception, Prediction, and the Origins of Human Knowledge // [http://www.research.ed.ac.uk/portal/files/9873993/Expecting\\_the\\_World\\_ACfinalNov2012.pdf](http://www.research.ed.ac.uk/portal/files/9873993/Expecting_the_World_ACfinalNov2012.pdf)

Dennett 2013 – *Dennett D. C.* Expecting ourselves to expect: The Bayesian brain as a projector // *Behavioral and Brain Sciences*. 2013. Vol. 36. No. 3. P. 209–210.

Fletcher, Frith 2009 – *Fletcher P. C., Frith C. D.* Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia // *Nature Reviews Neuroscience*. 2009. Vol. 10. No. 1. P. 48–58.

Fodor 1981 – *Fodor J. A.* Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology // *Fodor J. A.* RePresentations: Philosophical Essays on the Foundations of Cognitive Science. Cambridge, Massachusetts: A Bradford Book/The MIT Press, 1981. P. 225–253.

Gregory 1978 – *Gregory R. L.* Eye and Brain: The Psychology of Seeing. McGraw-Hill: New York, 1978.

Gregory 1980 – *Gregory R. L.* Perceptions as Hypotheses // *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 1980. Vol. 290. No. 1038. P. 181–197.

- Gregory 1997 – *Gregory R. L.* Knowledge in perception and illusion // Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences. 1997. Vol. 352. No. 1358. P. 1121–1128.
- Helmholtz 1962 – *Helmholtz H. von.* Concerning the Perceptions in General // Helmholtz H. von. Treatise on physiological optics. Vol. 3. Ch. 26. 3rd edn. Translated by J.P.C. Southall, 1925, *Op. Soc. Am.* Section 26, reprinted New York: Dover, 1962. P. 1–37.
- Hohwy 2013 – *Hohwy J.* The Predictive Mind. New York: Oxford University Press, 2013.
- Hohwy 2014 web – *Hohwy J.* The self-evidencing brain // <http://philpapers.org/archive/HOHTSB.pdf>
- Jeannerod 2006 – *Jeannerod M.* Motor Cognition: What Actions Tell the Self. New York: Oxford University Press, 2006.
- Kanizsa 1985 – *Kanizsa G.* Seeing and thinking // *Acta Psychologica*. 1985. Vol. 59. No. 1. P. 23–33.
- Kawato, Hayakawa, Inui 1993 – *Kawato M., Hayakawa H., Inui T.* A forward-inverse optics model of reciprocal connections between visual cortical areas // *Network: Computation in Neural Systems*. 1993. Vol. 4. No. 4. P. 415–422.
- Lee, Mumford 2003 – *Lee T. S., Mumford D.* Hierarchical Bayesian inference in the visual cortex // *Journal of the Optical Society of America*. 2003. Vol. 20. No. 7. P. 1334–1348.
- Mace 1977 web – *Mace W. M.* James J. Gibson's Strategy for Perceiving: Ask Not What's Inside Your Head, but What Your Head's Inside of // [http://www.trincoll.edu/~wmace/publications/Ask\\_inside.pdf](http://www.trincoll.edu/~wmace/publications/Ask_inside.pdf)
- Mumford 1992 – *Mumford D.* On the computational architecture of the neocortex: II the role of cortico-cortical loops // *Biological Cybernetics*. 1992. Vol. 66. No. 3. P. 241–251.
- Neisser 2014 – *Neisser U.* Cognitive Psychology. New York and London: Taylor and Francis, 2014.
- Noe 2004 – *Noe A.* Action in Perception. Cambridge, Massachusetts, London, England: The MIT Press, 2004.
- O'Regan, Noe 2001 – *O'Regan J. K., Noe A.* A sensorimotor account of vision and visual consciousness // *Behavioral and Brain Sciences*. 2001. Vol. 24. No. 5. P. 939–973.
- Rao, Ballard 1999 – *Rao R. P. N., Ballard D. H.* Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects // *Nature neuroscience*. 1999. Vol. 2. No. 1. P. 79–87.
- Rescorla 2015 web – *Rescorla M.* Bayesian Perceptual Psychology // <http://www.philosophy.ucsb.edu/docs/faculty/papers/bayesian.pdf>
- Rescorla 2016 web – *Rescorla M.* Bayesian Sensorimotor Psychology // [http://www.philosophy.ucsb.edu/docs/faculty/michael-rescorla/rescorla\\_bayesian-sensorimotor-psychology.pdf](http://www.philosophy.ucsb.edu/docs/faculty/michael-rescorla/rescorla_bayesian-sensorimotor-psychology.pdf)
- Rock 1983 – *Rock I.* The Logic of Perception. Cambridge, Massachusetts: The MIT Press, 1983.
- Seth 2015 web – *Seth A.* The Cybernetic Bayesian Brain // <http://open-mind.net/papers/the-cybernetic-bayesian-brain>
- Bernstein N. A.* Biomechanics and the Physiology of Movements / Ed. V. P. Zinchenko. Moscow, Voronezh: Izdatel'stvo «Institut prakticheskoi psikhologii», NPO «MODEK» 1997 (in Russian).
- Bruner J.* On Perceptual Readiness // *Psychological Review*. 1957. Vol. 64. No. 2. P. 123–152 (Russian Translation 1977).
- Kant I.* Kritik der reinen Vernunft. Hamburg: Felix Meiner, 1781 (Russian Translation 1994).
- Lektorsky V. A.* Realism, antirealism, constructivism and constructive realism in philosophy and science // *Constructivist Approach in Epistemology and Human Sciences*. Ed. V. A. Lektorsky. Moscow: Kanon+, 2009. P. 5–40 (in Russian).
- Marr D.* Vision. San Francisco: Freeman, 1982 (Russian Translation 1987).

*Neisser U.* Cognition and reality: Principles and implications of cognitive psychology. San Francisco: W. H. Freeman, 1976 (Russian Translation 1981).

*Sushchin M. A.* The Theory of Situated Cognition in Cognitive Science: A Critical Analysis. Candidate of Sciences in Philosophy Dissertation. Moscow, 2014 (in Russian).

*Sushchin M. A.* Are You Your Brain? Why Cognitive Neuroscience Needs Behavioral and Situated Approach // Arshinov V. I., Aseeva I.A., Budanov V.G., Grebenshchikova E.G., Grimov O.A., Kamenskii E.G., Moskalev I.E., Pirozhkova S.V., Sushchin M.A., Chekletsov V.V. Socio-Anthropological Aspects of Converging Technologies. Methodological Aspects / Ed. I. A. Aseeva, V. G. Budanov. Kursk: ZAO «Universitetskaya kniga», 2015. P. 191–210 (in Russian).

---

<sup>1</sup> Данное знаменитое правило было сформулировано Лапласом независимо от Байеса спустя полвека после публикации работы последнего.

<sup>2</sup> В целях упрощения изложения мы намеренно опускаем определенные технические детали, не имеющие для нас здесь принципиального значения – например, то, что мозг, как постулируется байесовскими моделями, кодирует информацию не в форме единичных оценок, а скорее распределений вероятности положения дел, или же что на окончательное принятие решений когнитивной системой могут влиять даже не апостериорные распределения вероятности сами по себе, а так называемая функция полезности, которая предписывает выбор наиболее оптимального решения в зависимости от апостериорного распределения (например, выбор оптимальной формы лечения в зависимости от вероятности того или иного заболевания).