ПСИХОЛОГИЧЕСКИЙ ЖУРНАЛ, 2004, том 25, № 2, с. 116–127

ЗА РУБЕЖОМ

МЕНТАЛИЗМ И БИХЕВИОРИЗМ: СЛИЯНИЕ?

© 2004 г. В. А. Лефевр*

Профессор Калифорнийского университета, Ирвин, США

Рефлексивно-интенациональная модель субъекта (RIMS) связывает биполярное вероятностное поведение субъекта с его ментальной сферой. Мы показываем, что закон соответствия (Matching Law) является формальным следствием этой связи. RIMS также позволяет теоретически вывести основные паттерны поведения животного в экспериментах с двумя альтернативами. Эта находка позволила нам выдвинуть гипотезу, что закон соответствия отражает процесс самопрограммирования субъекта, обладающего ментальной сферой. В результате субъект приобретает способность выбирать альтернативы с фиксированными вероятностями. Закон соответствия может служить операционным индикатором существования внутреннего мира у субъекта.

Ключевые слова: ментальный домен, паттерны поведения, внутренний мир субъекта, закон соответствия, утилитарная и деонтологическая оценка, саморефлексия.

ВВЕДЕНИЕ

Наука о субъективном, наделяющая живое существо "полостью" для внутреннего мира, и наука о поведении, отказывающая ему в этом, имеют одну общую черту: организм предстает в них в своей целостности. При этом в фокусе первой лежит отношение субъекта к самому себе, а в фокусе второй находится отношение субъекта и среды. В течение последних десятилетий демаркационная линия между ментализмом и бихевиоризмом существенно изменилась: была создана формальная модель субъекта, одновременно отражающая и его ментальный домен, и взаимодействие со средой. Верификация этой модели проявляется как ее экспансия в различные уже сложившиеся ветви психологии, социологии и антропологии. Однако именно бихевиоризм представляется наиболее привлекательным полем для такой экспансии. Причина лежит в его суровой внутренней дисциплине и методологической честности, позволяющих ясно отделять понятое от непонятого. Одной из главных нерешенных проблем науки о поведении является **закон соответствия** – ***Matching Law*** (Herrnstein, 1961), сущность которого заключается в способности птиц и млекопитающих *странно*, с точки зрения утилитарного здравого смысла, регулировать отношение между последовательностью реакций и последовательностью стимулов (см. Williams, 1988). В этой работе мы предлагаем решение этой проблемы с помощью рефлексивной модели интенционального субъекта (RIMS) (Lefebvre, 1992; 1999; 2001).

Создание этой модели стимулировалось попыткой понять, что собой представляет феномен "морального выбора", если подходить к нему не с моралистической, а с чисто научной точки зрения. Проблема вскрытия объективных законов морального выбора волнует огромное число специалистов, от психиатров до психологов, изучающих преступную и террористическую активность. В рамках таких исследований ментальный домен человека должен быть представлен столь же ясно и однозначно, как поведение в рамках бихевиоризма. RIMS – это особая математическая репрезентация субъекта, совершающего выбор между двумя альтернативами. Эта модель отражает два аспекта активности субъекта – утилитарный и деонтологический. Утилитарный аспект связан с практически выгодным поведением, направленным, например, на получение денег или пищи; деонтологический – с идеалистическим поведением, таким, например, как выбор между добром и злом. При этом "моральная" ориентация альтернатив может не совпадать с их утилитарной ориентацией. Например, сделка с врагом может оказаться практически более выгодной, чем договор с другом. Эти два различных аспекта связываются формализмом модели в единый процесс генерации поведения.

RIMS является вероятностной моделью. Она предсказывает вероятности, которые субъект выбирает альтернативы; одна из них играет для него роль позитивного полюса, другая – негативного. Мысль о том, что выбор субъекта имеет вероятностный характер, возникла в начале двадцатого века и была воплощена в нескольких теоре-

* Автор благодарен своим друзьями и коллегам В. Бауму, Д. Мазуру, Р. Касселю, В. Палея, Д. Райан, С. Шмидту и К. Виверу за ценные советы. Автор также хочет выразить благодарность Викторине Лефевр, без помощи которой эта работа не была бы сделана.

тических моделях (Thurstone, 1927; von Neuman, Morgenstern, 1944; Savage, 1951; Mosteller, Nogee, 1951; Bradley, Terry, 1952; Davidson, Suppes, Siegel, 1957; Bower, 1959; Luce, 1959; Audley, 1960; Spence, 1960; Restle, 1961; LaBerge, 1962; Atkinson et al., 1965). Эта линия исследований существенно изменила предшествующий взгляд на поведение как на процесс, однозначно детерминированный внешним миром. При этом, хотя были разработаны эффективные методы предсказания результатов вероятностного выбора, вопрос о его природе остался в стороне. У нас до сих пор нет ясного представления о том, все или только некоторые живые существа способны к вероятностному выбору, и как организм "узнает", с какими вероятностями он "должен" совершать выбор в данной ситуации. RIMS связывает вероятностное поведение субъекта с его ментальной сферой и позволяет сформулировать несколько новых гипотез. В рамках этой модели предполагается, что субъект перед актом выбора находится в неопределенном состоянии, которое может быть охарактеризовано распределением вероятностей выбора альтернатив. Используя квантово-механическую метафору, мы можем сказать, что субъект непосредственно перед актом выбора находится в смешанном состоянии, а сам акт выбора есть "коллапс" смешанного состояния, в результате чего субъект переходит в одно из чистых состояний. Следует подчеркнуть, что способность организма производить выбор альтернатив с *фиксированными* вероятностями говорит о его достаточно высоком уровне развития. Специалисты по математическому моделированию знают, как трудно создать техническое устройство, способное генерировать случайную последовательность нулей и единиц с заданной вероятностью их появления.

Мы можем предположить, что вероятностное поведение появляется одновременно с появлением у организма ментальной сферы. Их возникновение знаменует момент "освобождения" организма от "необходимости" однозначно реагировать на внешние воздействия. Чтобы выбирать альтернативы с определенными вероятностями, организм должен каким-то образом "загрузить" эти вероятности в самого себя. Мы полагаем, что "секрет" закона соответствия как раз и заключается в том, что он отражает процесс формирования у субъекта смешанного состояния, для перехода в которое субъект перерабатывает информацию, получаемую из внешнего мира в вероятностное распределение.

Представим себе, что организм голубя, крысы и даже человека не способен решить эту задачу посредством активности во внутреннем мире. Поэтому в процесс вычисления, по сути своей являющейся мыслительной активностью, вовлекается весь организм, и "беготня" животного в экспериментах с двумя кормушками (в которых обнаруживает себя закон соответствия) есть внешнее проявление этого процесса. В результате такой "загрузки" вероятности субъект оказывается способным произвести мгновенный вероятностный выбор. И эта способность достается субъекту не даром. Чтобы обладать ею, его организму приходится тратить энергию.

Эксперименты с двумя ключами, с человеком-испытуемым (см. Ruddle et al., 1979; Wearden, Burgess, 1982) позволяют предположить, что формирование смешанного состояния у людей также связано с моторной активностью. Она может проявлять себя и в других экспериментах. Например, при оценке интенсивности стимула с помощью категориальной шкалы карандаш испытуемого некоторое время колеблется, прежде чем делается окончательная отметка. Иногда даже трудно установить, какая отметка является финальной (Poulton, Simmonds, 1985). Мы можем предположить, что эти колебания являются функциональными аналогами беготни крысы между кормушками. Отметим, что RIMS объясняет и процесс категориальной оценки (Lefebvre, 1992). Нельзя исключить возможности, что у человека загрузка вероятностей может происходить и посредством движения глаз.

Наиболее важное отличие RIMS от прежде существовавших моделей заключается во введении в нее особой переменной, которая соответствует *модели себя* у субъекта (Lefebvre, 1965; 1992). Мы интерпретируем значение этой переменной как *интенцию* субъекта совершить выбор. Интенциональное поведение в RIMS задается условием $B = I$, где B – значение переменной, описывающей поведение субъекта, а I – значение переменной, соответствующей образу себя. В этом случае переменная I может быть опущена, и мы получаем модель бихевиористского типа, эмпирически тестируемую. В рамках RIMS организм субъекта стремится генерировать такую линию поведения, при котором достигается и поддерживается равенство $B = I$. Этот принцип генерации поведения мы далее называем *законом саморефлексии* (Lefebvre, 2002).

1. МЕСТО ЗАКОНА САМОРЕФЛЕКСИИ В ЛОГИЧЕСКИЙ СХЕМЕ ЭВОЛЮЦИИ БИХЕВИОРИЗМА

В эволюции науки о поведении прослеживается ясная логика, мало зависящая от индивидуальных предпочтений исследователей и от запретов на использование интроспективных понятий (как например, требование Павлова к помощникам не использовать выражений "собака заметила", "собака догадалась" и т.д.). Мы выделяем в развитии бихевиоризма четыре этапа и полагаем, что сегодня совершается переход к пятому (см. рис. 1.

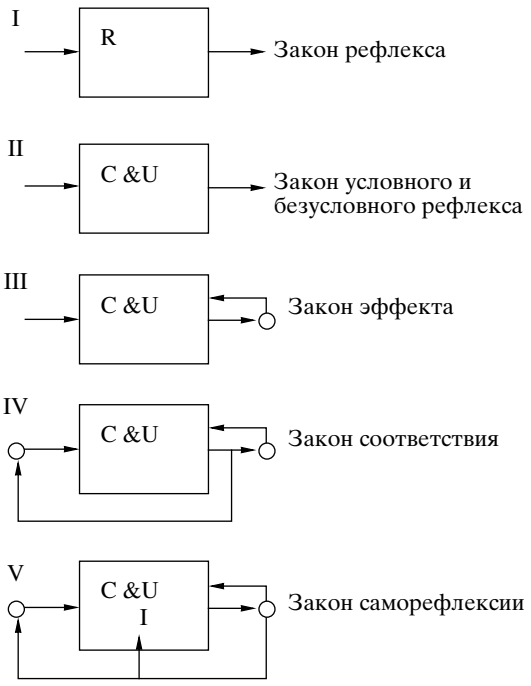

Рис. 1. Логические этапы развития бихевиоризма.

Схема не отражает временного порядка возникновения идей. Например, Павлов, Торндайк, Уотсон и Бехтерев работали примерно в одно время. Однако вклад раннего Уотсона и Бехтерева принадлежит первому этапу, Павлова – второму, а Торндайка – третьему.) Каждый этап может быть охарактеризован некоторым "законом", сжато выражающим определенное правило поведения живого организма.

На первом этапе, появившемся в эпохе картезианства, организм (если использовать более позднюю метафору) представлялся в виде черного ящика со входом и выходом. Выходу соответствуют акты жизнедеятельности организма, называемые реакциями, а входу – воздействия на субъекта со стороны окружающего мира. Внутри ящика находится некоторый механизм, *автоматически* трансформирующий каждое воздействие в реакцию. Такая трансформация была названа *рефлексом*.

Второй этап связан с открытием Павловым (1927) того факта, что рефлексы бывают двух типов: одни врожденные, а другие являются результатом индивидуального опыта организма. Первые получили название безусловных, а вторые – условных. Большинство вероятностных моделей поведения, в которых воздействию соответствует вероятностное распределение, принадлежат к первому или второму этапам. В отличие от детерминистских моделей они описывают статистические характеристики, а не функциональные отношения между стимулами и реакциями.

На третьем этапе было обнаружено, что автоматическая реакция на стимул может эволюционировать, становясь все более эффективной. Начало этого этапа было положено Торндайком (Thorndike, 1932), который сформулировал *закон эффекта*, отражающий способность живого существа модифицировать реакцию на стимул в зависимости от "эффекта" вызванного ее осуществлением. Например, организм кошки, запертой в клетке, производит селекцию успешных манипуляций с запором, и в конце серии экспериментов кошка, чтобы получить пищу, выходит из клетки быстрее, чем в начале (см. также Herrnstein, 1970; Williams, 1988).

Четвертый этап связан с новыми экспериментальными методами, разработанными Скиннером (Skinner, 1938) и его последователями. В опытах, которые они проводили, реакция животных могла влиять на устройство, генерирующее стимулы. Оказалось, что при таких условиях животное порождает особую линию поведения, в процессе течения которой между последовательностью стимулов и последовательностью реакций устанавливается устойчивое количественное соотношение. Математическая формула, выражающая это соотношение, была названа *законом соответствия*. Многочисленные попытки объяснить этот закон в рамках достаточно замкнутой системы понятий бихевиоризма не принесли убедительного успеха.

Взгляд на закон соответствия с точки зрения RIMS говорит нам о том, что этот закон есть проявление биполярности и закона саморефлексии (Lefebvre, 1999; 2002). Такие понятия как *образ себя и интенция* находятся за пределами словаря бихевиоризма. Поэтому без расширения концептуальных рамок наука о поведении рискует потерпеть поражение в попытках объяснить закон соответствия. Расширение же этих рамок будет означать переход к пятому этапу (рис. 1), который знаменует собой слияние ментализма и бихевиоризма.

2. ЗАКОН СООТВЕТСТВИЯ

Способность организма к регуляции связи между последовательностью реакций и последовательностью подкреплений была открыта Херренштейном (Herrnstein, 1961) в опытах с голубями. Стандартная камера для голубей была оборудована двумя ключами, клевки в которые могли вызвать подачу зернышка. Каждый ключ управлялся независимой программой, позволяющей варьировать средний интервал времени между подачами зернышек (программа VI – Variable-Interval). Эксперимент состоял из последовательных сессий, в каждой из которых средний интервал для каждого ключа был фиксирован. Пары интервалов подбирались так, чтобы иногда один

ключ выдавал подкрепления чаще, а иногда – другой.

Оказалось, что птицы выбирают такую линию поведения, при которой числа клевков в различные ключи (B_1 и B_2) примерно пропорциональны числам соответствующих подкреплений (r_1 и r_2):

$$\frac{B_2}{B_1} = \frac{r_2}{r_1}. \quad (2.1)$$

Равенство (2.1) получило название *закона соответствия*. Дальнейшие эксперименты проводились не только с голубями, но также с крысами и людьми. Помимо программ VI использовались и другие программы; например, VR (Variable-Ratio), при которых варьируется не временной интервал между последовательными подачами пищи, а среднее число клевков или нажатий на педали, необходимое для получения подкрепления. Результаты этих экспериментов привели к формулировке обобщенного закона соответствия (Baum, 1974):

$$\frac{B_2}{B_1} = c\left(\frac{r_2}{r_1}\right)^{\beta}, \quad (2.2)$$

где c и β-параметры, характеризующие поведение отдельного субъекта в эксперименте, состоящем из последовательности сессий. Сравнительно недавно Баум и др. (Baum et al., 1999) допустили возможность того, что закон (2.2) может быть сведен к соотношению

$$\frac{B_P}{B_N} = c\left(\frac{r_P}{r_N}\right), \quad (2.3)$$

где $B_P > B_N$. Величины с индексом P относятся к альтернативе, которую субъект выбирает чаще, а N – которую реже.

Соотношения (2.1), (2.2) и (2.3) являются частными случаями более общего соотношения

$$\frac{B_2}{B_1} = c\left(\frac{\varphi(r_2)}{\varphi(r_1)}\right), \quad (2.4)$$

отражающего поведение субъектов в описанных выше экспериментах (Davison, Jones, 1995; Baum, Aparicio, 1999).

3. ПОПЫТКИ ОБЪЯСНИТЬ ЗАКОН СООТВЕТСТВИЯ НЕ ВЫХОДЯ ЗА РАМКИ НАУКИ О ПОВЕДЕНИИ

Почему выполняется соотношение (2.4)? Естественно предположить, что оно является "побочным продуктом более фундаментальных процессов" (Williams, 1988). Доминирующая точка зрения на то, чем является этот процесс, содержится в следующих словах Баума и Апаричио: "Несмотря на ряд возражений, все ведущие теории относительно выбора животных в экспериментальной камере могут рассматриваться как модели оптимальности" (Baum Aparicio, 1999, с. 75). Идея оптимальности созвучна главному тезису философии бихевиоризма, в соответствии с которым животное так адаптируется к среде, что его поведение выглядит рациональным и целенаправленным.

Существует большая литература, в которой приводятся общие и экспериментальные аргументы *за* или *против* принципа оптимальности для объяснения закона соответствия (Williams, 1988; Baum et al., 1999). Весомым аргументом против являются результаты экспериментов, проведенных Мазуром (Mazur, 1981). В этих экспериментах голуби ставились в условия, при которых стремление к максимизации поступления пищи могло быть легко обнаружено. Птицы выбирали стратегии, не являющиеся оптимальными, однако такие, при которых выполняется соотношение (2.4).

Второй аргумент, который мы рассматриваем как критический для идеи оптимизации, требует более детального обсуждения. В выражении (2.2), описывающем обобщенный закон соответствия, есть два *свободных* параметра c и β. Значения этих параметров находятся экспериментально для каждого субъекта. Вопрос о причинах, заставляющих ввести параметр β, вызвал оживленную дискуссию (Baum, 1979; Wearden, Burgers, 1982; Aparicio, 2001), а параметр c чаще всего рассматривается просто как шкальная величина, связывающая полезности единиц подкрепления, получаемые от двух различных источников, и более подробно не обсуждается. Необходимость его введения обосновывается примерно такими рассуждениями. Представим себе, что кусочек пищи из левой кормушки составляет 0.75 от веса кусочка из правой. Если теперь записать соотношение (2.1) не для чисел кусочков (r_1 и r_2), а для суммарных "полезностей" кусочков, в данном случае равных весам, получаем:

$$\frac{B_2}{B_1} = \frac{0.75 r_2}{r_1}. \quad (3.1)$$

Подобного рода аргументация распространяется и на те эксперименты, в которых кусочки совершенно одинаковы. В этих случаях полагается, что параметр c отражает некоторый постоянно действующий, но скрытый фактор, изменяющий ценности единиц одного и того же продукта, полученного из двух различных источников. "Сдвиг (параметр c) – не нарушение животным принципа (закона соответствия), а неспособность экспериментатора контролировать все независимые переменные" (Baum, 1974, стр. 233). Иногда для спасения этой аргументации приходится допускать, что организм субъекта способен отобразить в

число c статистические характеристики наборов факторов, действующих неодновременно. Как иначе объяснить фиксированность c в экспериментах, где одна из альтернатив связана с программой VI, а другая с VR? (см., например, Baum, 1974, рис. 5). В этом случае трактовка константности c превращается в самостоятельную проблему, сравнимую по сложности с задачей объяснения закона соответствия. А если отказаться от интерпретации c как шкального коэффициента, то все известные нам попытки свести закон соответствия к максимизации оказываются неубедительными.

4. RIMS

Рефлексивно-интенциональная модель субъекта отражает внутренний мир и поведение субъекта, взаимодействующего с двумя различными объектами, которые далее будут называться "агентствами". Одно из агентств играет роль позитивного полюса, а другое – негативного. Субъект в RIMS представляется уравнением вида

$$X_1 - x_1 - (1 - x_1)(1 - x_2)M(x_3) = 0, \qquad (4.1)$$

где все переменные и функция $M(x_3)$ принимают значения из интервала [0, 1] (Lefebvre, 2001).

Переменная X_1 представляет исполнительную систему субъекта. Ее значение есть относительная вероятность, с которой исполнительная система *готова* воздействовать на позитивное агентство. Переменная x_3 соответствует модели себя у субъекта. Она представляет интенциональную сферу субъекта. Ее значение есть относительная вероятность, с которой субъект намерен воздействовать на позитивное агентство. При этом, возникновение *намерения* и возникновение *готовности* рассматриваются как два независимых события.

Переменная x_1 представляет характер отношений между субъектом и позитивным агентством в данный момент. Ее значение может интерпретироваться двумя способами: во-первых, как относительная частота, с которой позитивное агентство *воздействует* на субъекта; во-вторых, как *потребность* субъекта в воздействии на него позитивного агентства с такой частотой.

Переменная x_2 представляет опыт субъекта. Ее значение есть интегральная оценка относительной частоты, с которой позитивное агентство воздействовало на субъекта в прошлом.

Функция $M(x_3)$ представляет *прогностическую* активность субъекта. Ее значение есть субъективная оценка степени позитивности будущего, при условии, что интенция x_3 претворяется в реальность.

Среда может детерминировать все или только некоторые из значений X_1, x_1, x_2, x_3. Если набор значений, детерминированных средой таков, что ограничение (4.1) не реализуется, мы считаем, что субъект *дезадаптирован*. В противном случае, субъект *адаптирован* к среде и переменные, не детерминированные средой, могут принимать любые значения, не нарушающие (4.1).

Если значение x_1 детерминировано средой, оно интерпретируется как воздействие, полученное от среды; в противном случае, оно интерпретируется как потребность в воздействии со стороны среды.

Адаптированный к среде субъект называется *интенциональным*, если (4.1) дополнено ограничением

$$X_1 = x_3, \qquad (4.2)$$

т.е. у интенционального субъекта готовность к действию соответствует его интенции действовать.

Выбор интенционального субъекта мы называем *интенциональным вероятностным выбором*.

При условии (4.2) выражение (4.1) превращается в уравнение относительно X_1:

$$X_1 - x_1 - (1 - x_1)(1 - x_2)M(X_1) = 0. \qquad (4.3)$$

Отсутствие решений у этого уравнения означает, что субъект неспособен к интенциональному действию. В случае, когда это уравнение имеет решение, т.е. существует функция

$$X_1 = f(x_1, x_2), \qquad (4.4)$$

удовлетворяющая (4.3), мы можем исключить x_3 из рассмотрения, и RIMS превращается в бихевиористскую модель, все переменные которой могут быть соотнесены с наблюдаемыми величинами.

В случае, когда X_1 не есть функция переменных x_1 и x_2, мы полагаем, что субъект обладает способностью к выбору, даже вероятность которого не может быть предсказана внешним наблюдателями. Такой выбор мы называем *свободным*.

При моделировании конкретных психологических процессов могут вводиться дополнительные функциональные ограничения на связи между переменными X_1, x_1, x_2, x_3, отражающие специфические свойства моделируемых субъектов. Из уравнения (4.1) следует, что, независимо от вида функции $M(x_3)$, выполняются неравенства

$$x_1 \le X_1 \le 1 - x_2 + x_1x_2. \qquad (4.5)$$

В рамках этой работы мы полагаем, что функция $M(x_3)$ является линейной и имеет вид

$$M(x_3) = (1 - d)x_3, \qquad (4.6)$$

где $d \in [0, 1]$. Значение d интерпретируется как *индекс депрессии*, понижающий прогнозируемую субъектом степень позитивности будущего. Например, когда этот индекс принимает максималь-

ное значение $d = 1$, $M(x_3) \equiv 0$, будущее для субъекта негативно, а при $d = 0$, $M(x_3) \equiv x_3$, т.е. степень позитивности будущего всегда равна величине интенции. При условии (4.6), равенство (4.3) превращается в уравнение относительно X_1:

$$X_1 = x_1 + (1 - x_1)(1 - x_2)(1 - d)X_1,$$

из которого следует, что при $x_1 + x_2 + d > 0$

$$X_1 = \frac{x_1}{1 - (1 - x_1)(1 - x_2)(1 - d)}, \qquad (4.7)$$

а в случае, когда $x_1 = x_2 = d = 0$, величина X_1 не является функцией x_1 и x_2 и, в соответствии с определением, субъект обладает способностью к свободному выбору.

При $x_1 > 0$ соотношение (4.7) может быть представлено в виде

$$\frac{1 - X_1}{X_1} = (1 - (1 - x_2)(1 - d))\left(\frac{1 - x_1}{x_1}\right). \qquad (4.8)$$

Свяжем теперь это равенство с числами воздействия субъекта на различные агентства и числами воздействия различных агентств на субъекта. Положим, что

$$X_1 = \frac{N_1}{N_1 + N_2}, \quad x_1 = \frac{n_1}{n_1 + n_2}, \qquad (4.8a)$$

где N_1 и N_2 числа воздействий субъекта на позитивное и негативное агентства, а n_1 и n_2 числа воздействий, получаемых субъектом от позитивного и негативного агентств. Подставляя эти значения в (4.8), получаем

$$\frac{N_2}{N_1} = p\left(\frac{n_2}{n_1}\right), \qquad (4.9)$$

где

$$p = 1 - (1 - x_2)(1 - d). \qquad (4.10)$$

Выражение (4.9) есть одна из форм представления *интенционального субъекта* при $M(x_3) = (1 - d)x_3$. Со своей стороны, (4.9) соответствует обобщенному закону соответствия (2.2) при $\beta = 1$. Мы видим, что если (2.2) записано так, что B_1 соответствует позитивному полюсу, а B_2 – негативному, то свободному параметру c соответствует значение p.

Назовем величины

$$\frac{n_1}{N_1} = D_1 \quad \text{и} \quad \frac{n_2}{N_2} = D_2 \qquad (4.10a)$$

плотностями подкрепления D_1 и D_2. Теперь (4.9) можно записать как

$$\frac{D_1}{D_2} = p. \qquad (4.11)$$

Из эквивалентности (4.11), (4.9), (4.8) и (4.7) следует, что субъект интенционален тогда и только тогда, когда выполняется (4.11).

5. МОДЕЛИРОВАНИЕ ЭКСПЕРИМЕНТА С ДВУМЯ КОРМУШКАМИ

Мы полагаем, что активность субъекта в эксперименте с двумя ключами выполняет не только функцию добычи пищи, но и функцию генерации смешанного состояния (см. Введение). Сначала организм стабилизирует относительные частоты контактов с агентствами, поддерживая равенство $X_1 = x_3$. После стабилизации частота $N_1/(N_1 + N_2)$ "превращается" в равную ей *вероятность*, характеризующую смешанное состояние субъекта. Далее мы показываем, как эксперимент с двумя кормушками может быть промоделирован с помощью RIMS.

Субъект находится в клетке с двумя устройствами для подачи пищи, каждое связано с ключом, в который "клюет" субъект. Клевки достаточно редко подкрепляются выдачей кусочков пищи. Каждый ключ управляется независимой программой типа VI или VR. Эксперимент состоит из последовательных сессий. В каждой сессии средние интервалы подачи пищи фиксированы для каждого ключа. Положим, что

(1) Подготовка испытуемого к эксперименту (например, ограничение пищевого рациона) и условия его содержания в процессе эксперимента детерминируют величину индекса депрессии d, который сохраняется постоянным на протяжении всего эксперимента.

(2) В начале каждой сессии происходят следующие события:

(а) одна кормушка приобретает статус позитивного агентства, а другая – негативного;

(б) переменная x_2 принимает фиксированное значение, зависящее от средней частоты подкрепления кормушки, соответствующей позитивному агентству в предыдущих сессиях. Если данная сессия первая, то $x_2 = 1/2$.

Целью этого процесса является формирование и поддержание состояния, в котором у организма есть способность к мгновенному интенциональному вероятностному выбору. RIMS не говорит нам, какую стратегию выберет субъект, чтобы достичь и поддерживать соотношение (4.9). Возможно, субъект стремится сохранять соотношение (4.11) для локальных плотностей способом, похожим на тот, который дает модель мелиорации (Rachlin, 1973; Vaughan, 1985).

6. ПАТТЕРНЫ ПОВЕДЕНИЯ, КОТОРЫЕ ПРЕДСКАЗЫВАЕТ RIMS

Условимся вслед за Баумом и др. (Baum et al., 1999) называть предпочитаемой альтернативой (ключом) ту, к которой испытуемый обращается чаще в данной сессии. В рамках RIMS ключи поляризованы. Один является позитивным полюсом, а другой – негативным. Рассмотрим три возможных отношения между поляризацией и предпочтением на множестве сессий.

(A) Один из ключей является позитивным полюсом во всех сессиях, независимо от того, предпочтителен он или нет.

(B) В каждой сессии позитивным полюсом является непредпочитаемый ключ.

(C) В каждой сессии позитивным полюсом является *предпочитаемый ключ*.

Не теряя общности рассуждения, назовем один ключ правым, а другой левым. Пусть K_1 и K_2 числа клевков в правый и левый ключи, а k_1 и k_2 числа соответствующих подкреплений. Построим теперь графики величины $K_1/(K_1 + K_2)$ в зависимости от $k_1/(k_1 + k_2)$ и величины $\log(K_2/K_1)$ в зависимости от $\log(k_2/k_1)$ для случаев (A), (B) и (C). Величины, относящиеся к позитивному полюсу, мы будем обозначать по-прежнему N_1 и n_1, а относящиеся к негативному – N_2 и n_2. Для построения графиков используются функции (4.7) и (4.9). Каждый график относится к множеству сессия (см. рис. 2).

Графики A1 и A2 отражают случай, когда позитивным полюсом на всем множестве сессий является правый ключ, а негативным – левый.

Графики B1 и B2 соответствуют случаю, когда правый или левый ключи соответствуют позитивному полюсу только в тех сессиях, в которых они *непредпочитаемы*. Поэтому графики претерпевают разрыв. Рассмотрим B1. Для тех сессий, в которых $K_1 < K_2$, позитивным полюсом является правый ключ. В точке $K_1 = K_2$ происходит разрыв, соответствующий переориентации полюсов. При $K_1 > K_2$ позитивным полюсом является левый ключ. Логарифмический график (B2) состоит из двух лучей, идущих под углом 45° к горизонтальной оси. Левый луч соответствует сессиям, в которых позитивным полюсом является левая альтернатива, а правый – в которых правая.

Графики C1 и C2 отражают случай, когда правый или левый ключи соответствуют позитивному полюсу лишь в тех сессиях, в которых они *предпочитаемы*. Рассмотрим C1. Для тех сессий, в которых $K_1 > K_2$, позитивным полюсом является правый ключ. В точке $K_1 = K_2$, как и в случае (B), происходит разрыв. При $K_1 < K_2$ позитивным полюсом является левый ключ. Логарифмический график C2, так же как и график B2, состоит из двух лучей. Верхний луч соответствует случаю, когда позитивна левая альтернатива, а нижний – когда правая. Обратим внимание на различие между этими графиками: при $K_1 > K_2$ в B2 луч лежит выше диагонали, а в C2 – ниже; при $K_1 < K_2$ в B2 луч лежит ниже диагонали, а в C2 – выше. Сдвиг прямых вверх или вниз на логарифмических графиках A2, B2 и C2 предопределяется величиной

$$p = 1 - (1 - x_2)(1 - d).$$

Легко видеть, что $p = 1$ лишь при условии, что по крайней мере одна из величин (x_2 или d) равна 1. Значение $x_2 = 1$ означает, что все предшествующие подкрепления субъект получал от позитивного ключа. В реальных экспериментах у испытуемого всегда есть некоторый опыт получения подкрепления от негативного ключа, поэтому мы должны положить, что $x_1 < 1$.

Таким образом, идеальное соответствие

$$\frac{N_2}{N_1} = \frac{n_2}{n_1} \qquad (6.1)$$

может возникнуть лишь при $d = 1$, т.е. при условии, когда испытуемый приведен в состояние с *максимальным* индексом депрессии. В случае, когда $d = 0$, т.е. индекс депрессии принимает минимальное значение, реализуется соответствие

$$\frac{N_2}{N_1} = x_2\frac{n_2}{n_1}. \qquad (6.2)$$

7. НАБЛЮДАЕМЫЕ ПАТТЕРНЫ

Паттерн (A) хорошо известен. Его обычно описывают как случай, когда в равенстве (2.2) $\beta = 1$. Такой паттерн появляется при условии, что левая и правая альтернативы имеют некоторое существенное различие. Например, левый ключ управляется программой VI, а правый – программой VR (см. Baum, 1974; Williams, 1988). В этом случае множество сессий может описываться выражением (7.1), где $c \le 1$

$$\frac{B_1}{B_1 + B_2} = \frac{r_1}{r_1 + cr_2}. \qquad (7.1)$$

Этому уравнению соответствуют экспериментальные кривые, типа показанных на рис. 3.

В рамках RIMS этот паттерн может быть однозначно интерпретирован: ключ, которому соответствует B_1, является позитивным полюсом и $c = p$. Таким образом, (7.1) можно записать как

$$\frac{N_1}{N_1 + N_2} = \frac{n_1}{n_1 + pn_2}. \qquad (7.2)$$

Анализ экспериментов, в которых один ключ (скажем, левый) управляется программой VI, а

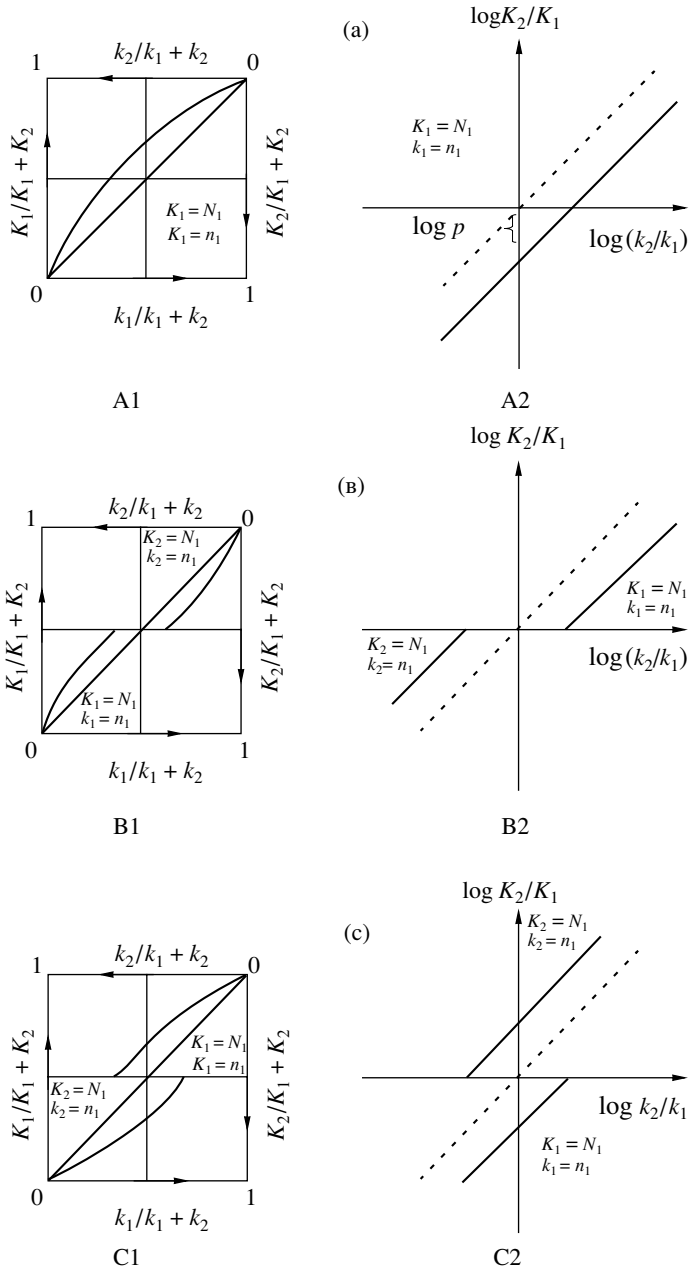

Рис. 2. Паттерны поведения, предсказываемые RIMS.

другой (правый) – программой VR, показывает, что ключ, связанный с VR, является позитивным полюсом, а связанный с VI – негативным.

Паттерн (B) также известен. Он наблюдается в тех случаях, когда альтернативы фактически ничем не отличаются кроме средних интервалов между подкреплениями. Это наблюдение склонило Баума и др. (Baum et al., 1999) предложить вместо обобщенного закона соответствия (2.2) – закон (2.3) (рис. 4).

Этот паттерн, как отметили Баум и др., соответствует случаям, для которых $\beta < 1$, если описывать их с помощью обобщенного закона соответствия. Появление характерного для этого закона изгиба кривой может, как указывают Баум и др., быть объяснено приближением разорванного графика 4(a) непрерывной степенной функцией (рис. 5).

Интерпретируя графики на рис. 4 с помощью RIMS, мы видим, что позитивному полюсу соответствует *менее* подкрепляемая альтернатива. Этот удивительный факт, как мы покажем ниже, является ключевым для понимания различия

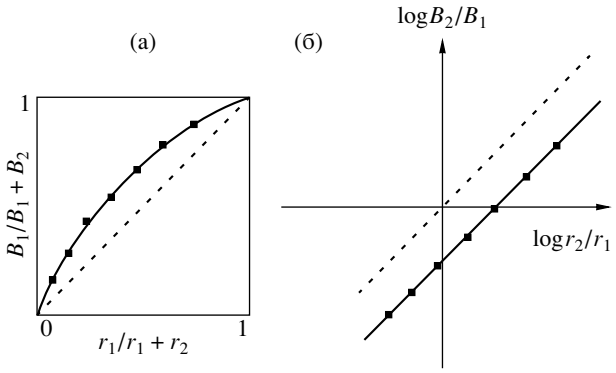

Рис. 3. Тип экспериментальных графиков, соответствующих паттерну (А).

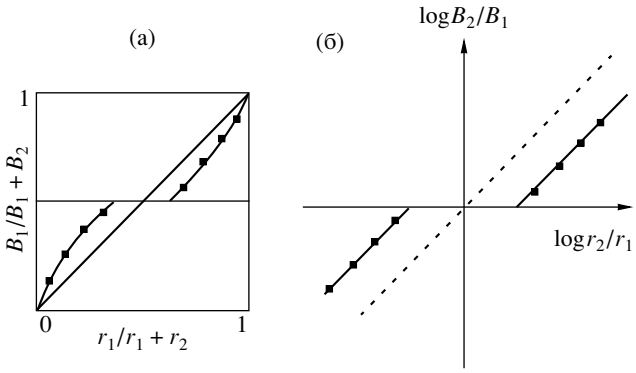

Рис. 4. Тип экспериментальных графиков, соответствующих паттерну (В).

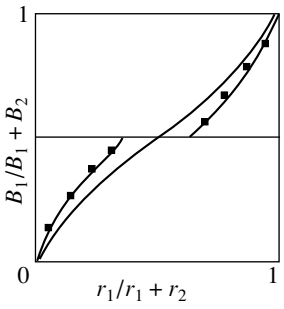

Рис. 5. Приближение разорванного графика (рис. 4) непрерывной степенной функцией.

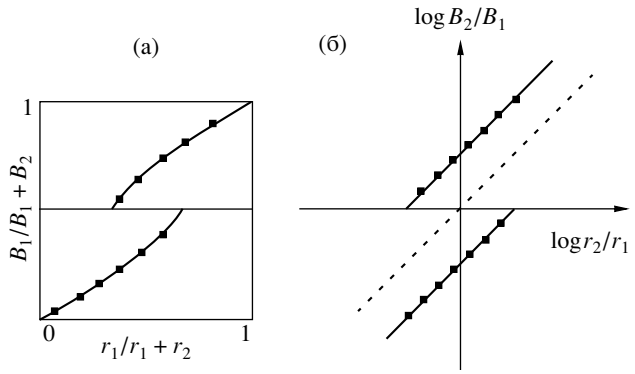

Рис. 6. Тип экспериментальных графиков, соответствующих паттерну (С).

между утилитарными и деонтологическими аспектами поведения животных.

Паттерн (С) встречается достаточно редко и как особый до сих пор не выделен. В качестве примера мы можем указать на эксперимент Баума и Апаричио (Baum, Aparicio, 1999), в котором одна из альтернатив была связана с программой VR с постоянным отношением, а вторая – с программой VI, интервалы которой менялись от сессии к сессии. Данные, полученные в этом эксперименте с крысами под номерами 102, 111, 120 и 213, могут быть представлены в виде следующего графика (рис. 6).

Можно допустить, что этот паттерн проявляет себя наиболее часто в тех экспериментах, результаты которых, если описывать их с помощью обобщенного закона соответствия, требуют введения $\beta > 0$. Как и в случае (В) появление изгиба кривой может быть объяснено приближением графика 6а степенной функцией (рис. 7).

Эксперимент Баума и Апаричио и анализ, проведенный ими (Baum, Aparicio, 1999), показывают, что паттерн (С) может быть сведен к паттерну (А), если альтернативы отождествлять не с пространственной локализацией (вправо–влево), а с типом программы (VR/VI). Тогда график 6а примет иной вид (рис. 8).

Мы видим, что паттерн (С) может появляться в тех экспериментах, в которых фактор, предопределяющий позитивно-негативную поляризацию альтернатив, в одних сессиях связан с левой альтернативой, а в других – с правой.

Каковы экспериментальные значения c в выражении (7.1)? Легче всего их найти для паттернов А. Для этого достаточно определить точку пересечения логарифмического графика с вертикальной осью на рисунке 3b. Как указывает Вильямс (Williams, 1988), в экспериментах, в которых один ключ управляется программой VI, а второй программой VR, величина $c = 0.59$. Для паттернов B большинство данных было обработано с предположением, что выполняется обобщенный закон соответствия (2.2), поэтому извлечь из них значение c практически невозможно. Чтобы все-таки найти значение с для паттерна B, пользуясь логарифмической формой представления данных, нужно отдельно осуществить линейную аппроксимацию для точек, лежащих ввыше горизонтальной оси, и для точек, лежащих ниже. Такая

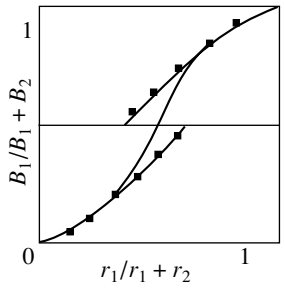

Рис. 7. Приближение разорванного графика – рис. 6а – непрерывной степенной функцией.

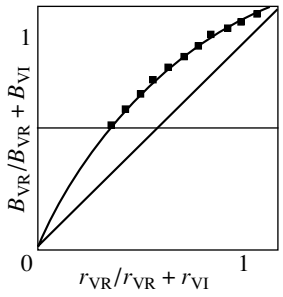

Рис. 8. Паттерн (A), соответствующий паттерну (C) на рис. 6а.

процедура по данным эксперимента с четырьмя голубями была проведена Баумом и др. (Baum et al., 1999). Среднее значение c, найденное нами по их данным, равно 0.58. В RIMS величине c соответствует величина p, даваемая равенством (4.10). Мы можем теперь найти среднее значение индекса депрессии d для этого эксперимента, положив $x_2 = 0.5$. Выражение (4.10) в этом случае имеет вид

$$0.59 = 1 - \left(1 - \frac{1}{2}\right)(1 - d), \qquad (7.3)$$

из которого находим, что $d = 0.18$. Следовательно, RIMS предсказывает, что прогностическая модель будущего у голубя в этом эксперименте представима в виде функции $m(x) = 0.82x$.

8. ДЕОНТОЛОГИЧЕСКИЕ ОЦЕНКИ У ЖИВОТНЫХ

Проведенное рассмотрение склоняет нас к выдвижению гипотезы, что у животных есть способность к биполярным оценкам типа “позитивно–негативно”, аналогичная (и, возможно, эволюционно предшествующая) моральным оценкам “хорошо–плохо” у человека. Как морально подобные биполярные оценки связаны с утилитарными предпочтениями, которые ясно проявляют себя в поведении животных? Чтобы приступить к ответу на этот вопрос, мы должны начать с анализа самих себя.

Рассмотрим один конкретный случай. В начале 20-х гг. прошлого века, во время гражданской войны в России, глава одной семьи попадает в Китай; его жену, оставшуюся с шестью детьми в России, расстреливают. Отец встречает богатого американца, который с риском для своей жизни едет в Россию, находит детей и через некоторое время привозит всех шестерых к отцу в Китай. Затем с огромными трудностями он помогает им переселиться в США. Старшая из спасенных детей, попав в США, испытывает глубокое разочарование.

Чем нехороша Америка была для двадцатидвухлетней Мули? Ностальгией? Да нет. А тем, что абсолютная для ценность – самопожертвование – приходит в столкновение с американским образом жизни. Муля растерялась: привычная ей идея “жизни для других” оказалась как-то неприложима к Америке. Слов нет, мистер Крейн “жертвует” много денег и времени на благотворительность (он помогает встать на ноги еще пятидесяти семьям!), но разве есть в его деятельности та особая красота тотального самопожертвования, восхищение которой с молоком матери впитала старшая дочь?! (Pann, 2003).

Давайте подавим в себе удивление (а, возможно, и возмущение) психологической неблагодарностью этой молодой женщины, попытаемся холодным взглядом натуралиста взглянуть на этот пример и понять, какая именно черта в поведении мистера Крейна оказывается неприемлемой для нее. Этот исключительно благородный человек оказал помощь пятидесяти одной семье. Совершенно ясно, что он должен был планировать свою деятельность и *считать деньги*, которые тратит на каждую семью. Другими словами, добро, которое он творил, оказалось соединенным с деньгами, и поэтому в его действиях отсутствует “особая красота тотального самопожертвования”. Создается впечатление, что в эту молодую женщину встроен автоматический механизм, препятствующий соединению утилитарных и деонтологических оценок.

Сделаем теперь следующий шаг. Допустим, что этот механизм лишь по своей форме выглядит культурно обусловленным, а по существу имеет глубокую биологическую природу. Развивая эту мысль, мы можем предположить, что у птиц и млекопитающих есть две принципиально различные системы оценок. Первая, *утилитарная*, отражает лишь оценку полезностей агентств, связанных с *ближайшими потребностями животного*; вторая, *деонтологическая*, связана с биполярными оценками “позитивно-негативно”, фиксирующими привлекательность агентств в более крупной шкале времени.

Рассмотрим, например, голодное животное, выбирающее одну из двух “кормушек”. Пусть первая кормушка более богата пищей, зато вторая потенциально более безопасна (скажем, хорошо укрыта от внешнего обзора). В этом случае более богатое пищей агентство получает оценку

"негативно", а менее богатое – "позитивно". Таким образом, "идеализм" животного связан с конкретными жизненно важными оценками, которые однако отделены от сиюминутных предпочтений. И нельзя исключить возможности того, что это различение поддерживается особым механизмом, работа которого у человека проявляется в драматическом противопоставлении материальных и идеальных ценностей.

Взглянем с этой точки зрения на паттерны поведения, описанные в предыдущих разделах. Начнем с паттерна (B). Он проявляется, когда левая и правая кормушки различаются только темпом подачи пищи. В этом случае, как мы установили, позитивным полюсом является *менее* подкрепляемая альтернатива. Мы можем допустить, что этот феномен отражает работу того же механизма, который у человека проявляется в противопоставлении "грязных" денег "чистым" помыслам. Поляризация альтернатив, противоположная предпочтению между ними как источниками питания, является аналогом человеческого акта *очищения*, отделения добра от практической пользы. Подчеркнем: это происходит лишь при условии, что альтернативы различаются только частотой подачи пищи. Если же существует некоторый не "утилитарный" фактор, предопределяющий поляризацию альтернатив, то поляризация сохраняется в течение всей сессии, причем независимо от того, является ли позитивная альтернатива более подкрепляемой или нет. Этот вывод есть результат интерпретации паттернов (A) и (C).

9. САКРАЛЬНЫЙ СДВИГ

Альтруизм, понимаемый как использование своих собственных ресурсов для помощи себе подобным, не является единственной формой жертвенного поведения. Добровольный расход средств и сил, связанный с созданием и поддержкой религиозных и моральных символов, является другой формой жертвенного поведения. Каждый из нас может привести примеры того, когда люди соглашаются выполнять работу, связанную с высшими ценностями, например, строить храм бесплатно или за меньшее вознаграждение, чем они требовали бы за эквивалентную работу, не связанную с такими ценностями.

RIMS предлагает объяснение этого феномена. Рассмотрим выражение (4.9). Мы можем интерпретировать числа N_1 и N_2 как *расходы* субъекта, которые он несет, обращаясь в позитивное и негативное агентства, а числа n_1 и n_2 как его доходы. Тогда

$$\frac{n_1}{N_1} \quad \text{и} \quad \frac{n_2}{N_2} \tag{9.0}$$

могут быть рассмотрены как средние платы, которые субъект требует от агентств за одно обращение к ним. Из (4.11) следует, что

$$\frac{n_1}{N_1} \le \frac{n_2}{N_2}. \tag{9.1}$$

Таким образом оказывается, что субъект в среднем никогда не берет за одно обращение в позитивное агентство больше благ, чем за одно обращение в негативное. И этот вывод справедлив не только для человека, но также для крыс и голубей.

ЗАКЛЮЧЕНИЕ

Мы уверены, что у нас есть внутренний мир. Эта уверенность основана только на нашем собственном субъективном опыте. Нет никакого *операционального критерия*, который бы позволил нам ответить на вопрос, имеет ли данный организм или техническое устройство субъективный мир, или же мы просто наблюдаем неодушевленные процессы. Эта проблема выходит далеко за рамки науки и касается основ нашей морали. Мы полагаем, что крыса или голубь могут страдать, но способны ли к страданию рыба или пчела?

В этой работе мы выдвигаем гипотезу, что внутренний мир живых существ появляется одновременно с их способностью совершать вероятностный выбор. Кроме того, мы обосновываем предположение, что существа с внутренним миром могут "самопрограммироваться". Другими словами, "загружать" в себя значения вероятностей, с которыми они будут делать свой выбор. Более того, мы показали, что закон соответствия есть внешнее проявление этого само-программирования.

Если эта гипотеза окажется правильной, закон соответствия будет служить операциональным критерием "одушевленности": мы сможем считать, что организмы, для которых он выполняется, имеют внутренний мир.

СПИСОК ЛИТЕРАТУРЫ

1. *Adams-Webber J.* Comment on Lefebvre's Model from the Perspective of Personal Construct Theory // J. of Social and Biological Structures. 1987. V. 10. P. 177–189.
2. *Adams-Webber J.* A Pragmatic Constructivist Gambit for Cognitive Scientists // Psycoloquy. 1995. V. 6(34).
3. *Adams-Webber J.* Self-reflexion in Evaluating Others // American J. of Psychology. 1997. V. 110. P. 527–541.
4. *Aparicio C.F.* Overmatching in Rats: The Barrier Choice Paradigm // J. of the Experimental Analysis of Behavior. 2001. V. 75. P. 93–106.
5. *Atkinson R.C., Bower G.H., Crothers E.J.* An Introduction to Mathematical Learning Theory. New York: Wiley, 1965.
6. *Baker H.D.* The Good Samaritan: An Exemplary Narrative of Moral Choice. Proceedings of the Workshop on Multi-Reflexive Models of Agent Behavior. Los Alamos, NM: Army Research Laboratory. 1999. P. 63–68.
7. *Batchelder W.H.* Some Critical Issues in Lefebvre's Framework for Ethical Cognition and Reflexion // J. of Social and Biological Structures. 1987. V. 10. P. 214–226.

8. *Baum W.M.* On Two Types of Deviation from the Matching Law: Bias and Undermatching // J. of the Experimental Analysis of Behavior. 1974. V. 22. P. 231–242.
9. *Baum W.M.* Matching, Undermatching, and Overmatching in Studies of Choice // J. of the Experimental Analysis of Behavior. 1979. V. 32. P. 269–281.
10. *Baum W.M., Aparicio C.F.* Optimality and Concurrent Variable-Interval and Variable-Ration Schedules // J. of the Experimental Analysis of Behavior. 1999. V. 71. P. 75–89.
11. *Baum W.M., Schwendiman J.W., Bell K.E.* Choice, Contingency Discrimination and Foraging Theory // J. of the Experimental Analysis of Behavior. 1999. V. 71. P. 355–373.
12. *Bradley R.A., Terry M.E.* Rank Analysis of Incomplete Block Design. The Method of Paired Comparisons // Biometrika. 1952. V. 39. P. 324–345.
13. *Davidson D., Suppes P., Siegel S.* Decision Making. Stanford: Stanford University Press, 1957.
14. *Davison M., Jones B.M.* A Quantitative Analysis of Extreme Choice // J. of the Experimental Analysis of Behavior. 1995. V. 64. P. 147–162.
15. *Herrnstein R.J.* Relative and Absolute Strength of Response as a Function of Frequency of Reinforcement // J. of the Experimental Analysis of Behavior. 1961. V. 4. P. 267–272.
16. *Herrnstein R.* On the Law of Effect // J. of the Experimental Analysis of Behavior. 191970. V. 13. P. 243–266.
17. *LaBerge D.L.* A Recruitment Theory of Simple Behavior // Psychometrika. 1962. V. 27. P. 375–396.
18. *Lefebvre V.A.* On Self-reflexive and Self-organizing Systems // Problemy Issledovania Sistem i Struktur. Moscow: Izdatelstvo AN USSR, 1965.
19. *Lefebvre V.A.* A Psychological Theory of Bipolarity and Reflexivity. Lewiston, N.Y.: The Edwin Mellen Press, 1992.
20. *Lefebvre V.A.* Categorization, Operant Matching and Moral Choice. Institute for Mathematical and Behavioral Sciences. MBS. 1999. 99-14, UCI.
21. *Lefebvre V.A.* Algebra of Conscience. Dordrecht: Kluwer Academic Publishers, 2001.
22. *Lefebvre V.A.* The Law of Self-Reflexion // J. of Reflexive Processes and Control. 2002. V. 1. P. 91–100.
23. *Luce R.D.* Individual Choice Behavior: A Theoretical Analysis. New York: Wiley, 1959.
24. *Mazur J.E.* Optimization Theory Fails to Predict Performance of Pigeons in a Two-Response Situation // Science. 1981. V. 214. P. 823–825.
25. *Mosteller F., Nogee P.* An Experimental Measurement of Utility // The J. of Political Economy. 1951. V. 59. P. 371–404.
26. *Pann L.* the Oldest Duaghter. Novoye Russkoye Slovo. 2003. May 24–25.
27. *Pavlov I.P.* Conditioned Reflexes. Oxford: Oxford University Press, 1927.
28. *Poulton E.S., Simmonds D.C.V.* Subjective Zeros, Subjectively Equal Stimulus Spacing, and Contraction Biases in Very First Judgments of Lightness // Perception & Psychophysics. 1985. V. 37. P. 420–428.
29. *Rachlin H.* Contrast and Matching // Psychological Review. 1973. V. 80. P. 217–234.
30. *Restle F.* Psychology of Judgment and Choice. New York: Wiley, 1961.
31. *Ruddle H., Bradshaw C.M., Szabadi E., Bevan P.* Behaviour of Humans in Concurrent Schedules Programmed on Spatially Separated Operanda // Quarterly J. of Experimental Psychology. 1979. V. 31. P. 509–517.
32. *Savage L.J.* The Theory of Statistical Decision // American Statistical Association J. 1951. V. 46. P. 55-67.
33. *Skinner B.F.* The Behavior of Organisms: An Experimental Analysis. New York: Appleton-Century-Crofts, 1938.
34. *Spence K.W.* Conceptual models of Spatial and Non-Spatial Selective Learning // Behavior Theory and Learning, Englewood Cliffs / Ed. Spence K.W. N.J.: Prentice-Hall, 1960.
35. *Thorndike E.L.* Animal Intelligence: An Experimental Study of the Associative Processes in Animals // Psychological Rview Monographs Supplement. 1911. V. 2. № 8.
36. *Thurstone L.L.* A Law of Comparative Judgment // Psychology Review. 1927. V. 34. P. 273–286.
37. *Vaughan W.* Choice: a Local Analysis // J. of the Experimental Analysis of Behavior. 1985. V. 43. P. 383–405.
38. *von Neuman J., Morgenstern O.* Theory of Games and Economic Behavior. Princeton: Princeton University Press, 1947.
39. *Wearden J.H., Burgess I.S.* Matching Since Baum (1979) // J. of the Experimental Analysis of Behavior. 1982. V. 28. P. 339–348.
40. *Williams B.A.* Reinforcement, Choice, and Response Strength. In: Atkinson, R.C., Herrnstein R.J., Lindzey, G., Luce, R.D. (Eds.) // Steven's Handbook of Experimental Psychology. 1988. V. 2. P. 167–244.

MENTALISM AND BEHAVIORISM: MERGING?

V. A. Lefebvre

Professor, University of California at Irvine, USA

The Reflexive-Intentional Model of the Subject (RIMS) connects the subject's bipolar probabilistic behavior with its mental domain. We demonstrate that the Matching Lawis a formal consequence of this tie. RIMS allows us also to deduce theoretically the main patterns of animal behavior in the experiments with two alternatives where the Matching Law reveals itself. This finding inclines us to put fortha hypothesis that this law reflects the process of self-programming of the subject with mental domain. As a result, the subject acquires the ability to choose alternatives with fixed probabilities. With this explanation, the relative frequencies of pressing a pedal or pecking at a key play the role of half-finished-products which after being downloaded into the self turn the probabilities of choice. The Matching Law can be regarded at as an operational indication of the mental domain existence.

Key words: mental domain, patterns of behavior, subject's inner world, matching law, utilitarian and deontological assessment, selfreflection.