Серия "ПРАКТИКУМ"
Выпуск 2

**А.Н. Гусев, Ч.А. Измайлов,
М.Б. Михалевская**

ИЗМЕРЕНИЕ
В ПСИХОЛОГИИ

Общий психологический практикум

Москва
«Смысл»
1987

ПРЕДИСЛОВИЕ

С момента выхода учебного пособия Ч.А. Измайлова, М.Б. Михалевской “Общий практикум по психологии. Измерение в психологии”, выпущенного небольшим ротапринтным тиражом в Издательстве Московского университета, прошло 13 лет. Произошли изменения не только в учебных планах, но и в техническом оснащении практикумов на факультетах психологии многих университетов. Появилась новая учебно-методическая литература, широкое распространение в учебном процессе получили персональные компьютеры. Настоящая книга представляет собой исправленное и значительно переработанное издание этого учебного пособия. Некоторые разделы исключены, другие написаны заново. Подготовлено много новых компьютерных учебных заданий.

Для психологии, как и для любой другой науки, процедуры измерения психологических переменных дают возможность устанавливать количественные связи между психологическими характеристиками и тем самым формулировать психологические законы. Кроме того, многие практические приложения психологии прямо основаны на проведении измерений. В этом смысле измерение служит главной силой, преобразующей психологию из науки описательной, следующей за фактами, в науку, умеющую предсказывать. Для студентов-психологов почти сразу же становится очевидной необходимость измерения в исследовании когнитивных процессов, где уже сформулирован целый ряд общих законов, но не менее важны измерения и при изучении эмоционально-волевой сферы психики, где также существуют количественные связи между различными психологическими характеристиками, примером чему может служить закон Йеркса—Додсона, связывающий успешность решения задачи с уровнем мотивации и со сложностью задачи.

Конечно, это вовсе не означает, что психологическое исследование исчерпывается измерением. Измерительная процедура — это только инструмент психолога, как, например, пила или рубанок — это инструмент столяра. Целью его деятельности является изготовление стула или шкафа, а не пиление или строгание само по себе, точно также целью психолога является решение с помощью измерений некоторой психологической задачи. Иначе говоря, измерение психологических переменных — необходимое, но не достаточное условие исследования. Но как нельзя стать столяром, не научившись профессионально оперировать различными столярными инструментами, точно так же нельзя стать профессиональным психологом, не научившись измерительным процедурам. Для этого необходимо, чтобы специалист-психолог не только профессионально владел широким набором измерительных процедур, существующих в психологии, но и сумел выбрать, а в случае необходимости и модифицировать стандартную измерительную процедуру адекватно решаемой задаче. Этим определяется и то большое значение, которое придается данному разделу в программе Общего психологического практикума, а также в разделе “Ощущение и восприятие” курса общей психологии.

Все методы психологического шкалирования, которые вошли в данную книгу, разделены на три класса: 1) *методы измерения чувствительности*; 2) *методы одномерного шкалирования*; 3) *методы многомерного шкалирования*. Эта классификация основывается на том принципе, что с точки зрения теории измерений все множество различных измерительных процедур, применяемых в психологии, является процедурами построения шкал психологической переменной (или процедурами психологического шкалирования) и различие между ними определяется двумя аспектами: 1) *типом построения шкалы*, т.е. является она шкалой наименований, шкалой порядка и т. д., либо 2) *степенью структурной сложности шкалы*, которая может иметь нулевую размерность, быть одномерной или многомерной.

Именно последний аспект положен в основу нашей классификации, поскольку она, на наш взгляд, имеет преимущество с точки зрения дидактики и позволяет показать студенту последовательное усложнение тех характеристик объекта, которые мы измеряем различными методами.

Здесь имеется в виду следующее. С помощью первого класса методов находится *одно-единственное значение на психологической шкале,* или, если характеризовать эти методы в геометрических терминах (как это часто делают), то можно сказать, что первый класс методов предназначен только для определения места точки в *психологическом пространстве*[1]. Поэтому методы измерения чувствительности можно назвать *методами локализации точки на психологической шкале.*

Второй класс методов позволяет определить всю психологическую шкалу в целом, но это *одна-единственная шкала,* характеризующая только одну простую психологическую характеристику. Другими словами, эти методы предназначены для определения *только одной из осей* психологического пространства. Поэтому их можно назвать методами построения одномерной психологической шкалы, или *методами одномерного шкалирования.*

И, наконец, третий класс методов предназначается для измерения сложных многомерных психологических характеристик. С их помощью строится целостная система шкал, определяющая взаимосвязь сразу нескольких психологических характеристик, иначе говоря, строится уже целостное психологическое пространство. Поэтому они названы методами построения сложных, многомерных шкал, или *методами многомерного шкалирования.*

В каждый из трех классов подбирались методы, имеющие наиболее важное значение с точки зрения профессиональной подготовки специалиста-психолога, т.е., во-первых, наиболее детально и глубоко разработанные как в теоретическом, так и в процедурном плане, во-вторых, наиболее широко применяющиеся в научно-исследовательских и прикладных работах, в-третьих, полностью исчерпывающие тот обязательный объем знаний, умений и на-

[1] Термин "психологическое пространство" будет часто употребляться на страницах этой книги. Здесь мы только подчеркнем, что в известной степени — это хорошая описательная метафора, имеющая явную содержательную аналогию с понятием геометрического пространства.

выков, который необходим для получения методической грамотности.

Внутри каждого класса методов их последовательность связана главным образом со сложностью математической модели шкалирования, лежащей в основе каждого метода измерения. Это сделано с дидактической целью, чтобы каждый шаг последовательно умножал опыт студента. В тех случаях, где это правило не нарушалось, в изложении методов сохранялась также и историческая последовательность их создания.

Измерительная процедура метода представляет собой алгоритм из набора определенных операций. Последовательность и взаимосвязь этих операций определяются теоретической моделью психологического шкалирования, в рамках котороой и разрабатывается тот или иной метод. Однако зачастую для обоснования одного и того же метода или его модификации предлагаются новые, иногда даже альтернативные модели шкалирования, с новой интерпретацией получаемых данных. Поэтому наряду с операциональным описанием методических процедур измерения, при изложении каждого метода рассматриваются относящиеся к нему наиболее распространенные теоретические модели шкалирования. Рассмотрение последних имеет большое значение еще и потому, что ими во многом определяется набор необходимых математических методов (а в последние годы и статистических программ) анализа результатов измерения. Наконец, особенно важно помнить о теоретической модели шкалирования в случаях, когда измерение психологических характеристик поставляет материал для построения психологической модели более высокого порядка; в этом случае особенно важно, чтобы не возникало скрытых противоречий между аксиоматикой одной и другой модели, приводящих к явным парадоксам в интерпретации результатов.

В работе над учебным пособием авторы получили большую помощь от своих коллег — сотрудников факультета психологии Московского государственного университета им. М.В. Ломоносова. Теоретическая часть главы “Методы обнаружения сигналов” написана бывшим сотрудником фа-

культета, профессором Э. Джафаровым. Большинство компьютерных программ подготовлены нашими коллегами Д. Аракельянцем, А. Кремлевым и Д. Чекалиным. Большую помощь в перепечатке рукописи оказала О. Малова. Оригинал-макет книги и часть рисунков подготовлены Д. Чекалиным. Всем им огромное спасибо.

ВВЕДЕНИЕ В ПСИХОЛОГИЧЕСКОЕ ШКАЛИРОВАНИЕ

§1. Психофизические шкалы

Самые первые методы психологических измерений (т.е. методы построения психологических или субъективных шкал) были разработаны в разделе психологии, называемом *психофизикой*. Основная задача, которую ставили перед собой психофизики — это определить, как соотносятся физические параметры стимуляции и соответствующие им субъективные оценки наших ощущений (Боринг, 1950). Зная эту связь, т.е. имея в распоряжении функцию типа $R=f(S)$, где S — значение физического параметра стимула, а R — значение субъективной реакции, предсказать ощущение, соответствующее какому-либо стимулу, есть дело простого расчета.

Психофизическая функция устанавливает связь между числовыми значениями двух типов, с одной стороны — это шкала физического измерения стимула, с другой — значения психологической или субъективной реакции на этот стимул. Очевидно, что точность расчетов прямо зависит от функции связи f, т.е. от того, будет она более жесткой или более расплывчатой. Но сама психофизическая функция, как шкала, в свою очередь, зависит от того, что из себя представляют исходные измерения R и S. Например, если измерения R и S дают шкалу отношений, то функция f может устанавливать пропорциональную зависимость, а если R и S являются только порядковой шкалой, то и результирующая связь между ними ограничится установлением монотонности, например, но не более. Таким образом, для построения психологических шкал существенно, какого типа измерение было проведено как для стимулов, так и для реакций. Но в то время как физические измерения достаточно хорошо известны и пользуются у исследователей доверием, психологические измерения даже в среде психологов популярны намного меньше, поэтому мы несколько подробнее отметим принципы измерений, относящихся к субъективному шкалированию.

В основе субъективных измерений лежит процедура приписывания чисел элементам из данного множества реакций. Это приписывание должно производиться по некоторым правилам. Правила заключаются в том, чтобы определенные *отношения*, которые установлены для чисел, выполнялись также и на множестве реакций. В зависимости от того, какие именно отношения можно установить для данного множества реакций, строится и соответствующая шкала измерения. По общепринятой классификации для субъективных измерений обычно рассматривают четыре основных типа шкал (Стивенс, 1960; Пфанцагль, 1976):

1. Шкала наименований, или классификационная шкала строится на единственном отношении — *отношении эквивалентности*. Деления на шкале характеризуют критерии, на основании которых производится классификация. Способность человека оценить любой стимул по заданному критерию как принадлежащий или не принадлежащийданному классу настолько очевидна, что возможность построения шкалы наименований для реакций различного уровня сложности обычно не вызывает возражений.

2. Шкала порядка строится на основании сразу двух отношений — эквивалентности и *порядка*. Естественно, что далеко не все объекты субъективно можно подчинить отношению порядка. Например, сразу очень трудно сказать, что больше — круг или треугольник, однако если выделить в этих объектах такое свойство, как площадь, то установить порядковые отношения для этих объектов уже значительно легче. Такие упорядочивания объектов по их отдельным свойствам широко используются при составлении различных оценочных шкал.

3. Шкала интервалов. Этот тип шкалы требует дополнительной возможности устанавливать равенство попарных различий между двумя парами стимулов, иначе говоря, определять *равенство субъективных интервалов*. Возможность построения такой шкалы позволяет большую часть свойств существующих числовых систем приписывать тем числам, которые получены на основе субъективных оценок. Постро-

ение для реакций шкалы интервалов является в психологии уже значительным достижением. Но, с другой стороны, интуитивно не очевидно, что человек всегда может делать оценки, соответствующие шкале интервалов. Действительно, если субъективные оценки не соответствуют некоторому физически измеряемому свойству, то совсем не очевидно, как можно установить соответствие оцениваемых стимулов шкале интервалов.

4. Шкала отношений получается, когда, кроме уже перечисленных операций: эквивалентности, порядка и сравнения разностей — можно осуществить для объектов *сравнение попарных отношений*. Это обусловлено возможностью оценивать абсолютное значение величины реакции и требует наличия на шкале нулевой точки, как на шкале температур Кельвина, например.

Последние две шкалы можно назвать "сильными", т.е. по результатам таких измерений можно строить более точные, более однозначные психофизические функции, к ним можно применить более тонкий статистический аппарат, чего нельзя сделать по отношению к первым двум типам шкал. Немаловажное значение имел и тот факт, что все физические измерения приводят обычно именно к "сильным" шкалам. Поэтому для психофизиков большее значение имели методические процедуры построения шкал интервалов и отношений, чем каких-либо других.

Первые две шкалы получили название *неметрических*, вторые две — *метрических*. В соответствии с этим в психологии говорят и о двух подходах к психологическим измерениям: метрическом (более строгом) и неметрическом (менее строгом).

§2. Нольмерное шкалирование

Во многих психологических исследованиях возникает задача определения единственного или специального значения психологической переменной, аналогично, например, задаче нахождения экстремума функции в математике. Такое специальное значение психологической переменной называется *порогом*. Впервые в психологии эта проблема была поставлена *Э. Гербертом*, как задача определения порога сознания — критической точки на континууме состояний от совершенно неосознанного до ясного сознания. Основной вклад в создание процедур пороговых измерений был сделан *Г. Фехнером* (1860), разработавшим первые методы пороговых измерений. Последующее развитие экспериментальной психологии показало, что порог является универсальной психологической характеристикой, и пороговые измерения получили широкое распространение, особенно в исследованиях познавательных процессов — восприятия, внимания, памяти.

В связи с их специфичностью *пороговые методы* обычно отделяют от остальных методов шкалирования психологических переменных (Вудвортс, Шлоссберг, 1958; Торгерсон, 1958). Однако это основание, разделяющее психологические измерительные процедуры на пороговые методы и методы шкалирования, является чисто содержательным, поэтому оно менее существенно, чем формальное основание, которое объединяет их вместе.

В терминах теории измерений определение порога есть нахождение одного шкального значения или локализация точки на психологической шкале. В формальном смысле — это построение психологической шкалы, имеющей единственное значение и нулевую размерность.

Поэтому все пороговые методы можно также рассматривать как методы построения психологических шкал, а развитие методов психологического шкалирования рассматривать, соответственно, как разработку процедур, позволяющих постепенно увеличивать размерность психологической шкалы. С этой точки зрения пороговые измерения являются самым простым видом психологического шкалирования. Следующий шаг в развитии психологических измерений состоял в разработке методов, позволяющих построить шкалу, содержащую все значения данной психологической переменной. Причем в качестве окончательного результата измерения стремились получить именно “сильную” шкалу.

§3. Одномерное шкалирование

Первый вклад в создание этих процедур был сделан также Фехнером (1860), разработавшим первую модель одномерного шкалирования, но основную детальную проработку процедур одномерного шкалирования осуществил *Терстоун* (1927, 1929), а затем *Стивенс* и его сотрудники (1937, 1955), разработавшие методы прямой оценки стимуляции. Далее эти методы развивались в работах шведских психофизиков под сильным влиянием *Экмана* (1965). Разработанные ими методы построения “сильных” шкал дали возможность психологам быстро продвинуться в решении многих психологических проблем, связанных с различными областями познавательных процессов. Эти методы стали широко распространяться, и здесь сразу же появились принципиальные ограничения, связанные с двумя особенностями этих методов: во-первых, с необходимостью выделения простой, одномерной психологической стимуляции, и, во-вторых, с наличием заранее известной физической шкалы измерения стимула. Но даже когда для стимула существует однозначная физическая шкала измерения, испытуемые, устанавливая метрические отношения между простыми субъективными реакциями, сталкиваются с трудностями. На это указывает большая вариабельность производимых испытуемым оценок. Зачастую она превосходит величину самой оценки в несколько раз (Пьерон, 1966).

Операции установления порядка или эквивалентности значительно проще и стабильнее. Существенным достоинством *порядкового шкалирования* является возможность его применения для измерений таких стимулов, которые в силу своей сложности не поддаются жестким, метрическим измерениям. Именно поэтому процедуры построения шкал порядка чрезвычайно распространены в таких разделах психологии, как психодиагностика, исследование эмоций, интеллекта и т.д.

Такие разные, но необходимые свойства измерительных процедур, как простота и стабильность “слабых” шкал, и точность “сильных” шкал, привели к идее создания такой процедуры, которая позволяет строить шкалу интервалов или шкалу отношений на основе оценок порядка или эквивалентности. Такие шкалы можно назвать *производными* шкалами интервалов или отношений в отличие от *первичных*, о которых шла речь выше. Для первичных шкал субъективные операции над объектами (их оценка или сравнение) и числовые операции связаны друг с другом непосредственно, без всякой промежуточной процедуры. Производная шкала методически имеет более сложную структуру, она строится с помощью дополнительной процедуры на базе первичной шкалы и, естественно, что такая процедура имеет смысл, только если производная шкала будет “сильнее” первичной. “Сила” производной шкалы основывается на теоретических допущениях о том, что исследуемые субъективные реакции обладают дополнительными свойствами кроме тех, которые установлены эмпирическими операциями, иначе говоря, здесь предусматривается построение развитой *модели шкалирования.*

Примером построения производной шкалы может служить модель шкалирования Фехнера. В основе модели лежат эмпирические процедуры, устанавливающие для стимулов отношение равенства и порядка. Например, в случае *метода “средней ошибки”* испытуемому предлагается, по сути дела, производить бинарную классификацию (ответы “да-нет”, “равны-неравны”), подравнивая переменный стимул к стандартному. При многократном повторении этой процедуры значение подравниваемого (переменного) стимула распределяется около значения стандартного в некотором диапазоне неразличимости. Вводится теоретическое предположение, что полученное таким образом распределение имеет форму *нормального распределения* и величина дисперсии этого распределения принимается за *меру порогового различия* переменного и стандартного стимулов на субъективной шкале. Далее делается *допущение равенства таких мер* во всех точках шкалы и, следовательно, вводится *единица измерения* на шкале; точка абсолютного порога принимается за нуль шкалы, и, таким образом, строится шкала отношений.

Конечно, эти допущения не для всякого случая справедливы, однако там, где их можно сделать, можно построить и стабильную шкалу отношений, основываясь на простых оценках.

Процедуры построения первичных и производных шкал позволили решить задачу построения точной психофизической функции в области простых ощущений, таких как видимая яркость, громкость различных звуковых тонов, тяжесть и т.п. Существуют достаточно надежные методы для получения экспериментальных данных, на основании которых определяются психофизические функции, связывающие субъективные шкалы с физическими. В случае наличия очевидного физического континуума стимулов, подобно интенсивности светового излучения при измерениях субъективной яркости света, несомненно, что этот континуум можно использовать как базу для построения точной субъективной шкалы. И как только психофизики показали, какими физическими свойствами объектов пользуются люди для своих оценок, соотносить физические шкалы с субъективными оценками стало достаточно простым и надежным делом, что и позволило психофизике из теоретических разделов перейти в прикладные.

§4. Модель шкалирования Терстоуна

Однако, по мере того, как исследователи переходили к ситуациям, все более приближенным к естественному поведению, становилось очевидным, что однозначно определить физический континуум для каждой важной субъективной характеристики невозможно. Это особенно ясно, когда речь идет об оценке достаточно сложных стимулов, таких как интеллектуальные способности или успеваемость в обучении.

Новое направление в психофизике возникло в связи с вопросами субъективного шкалирования характеристик стимула, *не имеющих однозначной физической интерпретации*. На основании *метода парных сравнений* (Кон, 1894) были разработаны процедуры для получения шкалы интервалов и отношений для таких стимулов. Формальным математическим основанием для этих процедур послужила *модель* *шкалирования сравнительных суждений* Терстоуна (1927, 1929, 1975).

Рассмотрим, к примеру, способ возможного создания шкалы “красоты” цвета с использованием закона сравнительных суждений. Стимулами в этом случае могут служить карты, окрашенные в различные цвета. Испытуемому предлагается рассмотреть по очереди все возможные пары карт и в каждой паре выбрать более “красивый” цвет. Каждая пара предъявляется много раз, и определяется *частота предпочтения* каждого цвета при сравнении его с остальными.

На этом первом шаге для исходного набора цветных карт строится порядковая шкала частот предпочтения. На втором шаге вводится теоретическое допущение, что дисперсия процесса различения каждой пары стимулов распределена по *нормальному закону*. Тогда субъективная разность (различие) между двумя стимулами S_i и S_j может быть измерена в единицах дисперсии, которая, в свою очередь, может быть оценена по наблюдаемой в опыте частоте суждения типа: стимул S_i более красивый, чем стимул S_j. Определение шкальных значений попарных разностей даст возможность построить шкалу интервалов для такого свойства стимулов, как красота цветных карт, хотя физические корреляты этой красоты остаются *неизвестными*.

Получив возможность расположить стимулы по субъективной шкале красоты, можно перейти к обратной процедуре — выявить, какой физический параметр в стимулах меняется в соответствии с полученной шкалой, и проверить, можно ли интерпретировать этот параметр как физический коррелят красоты.

§5. Многомерный анализ сложных стимулов

Однако модель Терстоуна предполагает обязательную одномерность шкалируемого свойства объектов, в данном случае “красоты” цветных карт, т.е., независимо от того, сколько физических характеристик стимула определяет оценку “красоты” цвета, психологически все карты должны быть выстроены в некоторую последовательность по степени красоты. Если же в действительности для оценки красоты цвета используется больше, чем одна субъективная размерность (как, например, при оценке различий между аппертурными цветами: одна — для цветового тона, другая — для насыщенности), то, используя модель Терстоуна, построить адекватную шкалу “красоты” цвета невозможно. В лучшем случае это будет какая-то проекция действительных шкал в одномерное пространство, на одну шкалу, и восстановить эти исходные шкалы по имеющейся единственной шкале, конечно, невозможно.

Естественно, что в этом случае невозможно решить и главную задачу, т.е. построить психофизическую функцию, поскольку невозможно обнаружить те физические свойства стимула, которые объясняют субъективные оценки.

Задачи построения сложных “многомерных” субъективных шкал и их последующей связи со шкалами физических свойств стимуляции породили целый ряд методов так называемого *многомерного анализа* (факторный анализ, многомерное шкалирование, дискриминантный и кластерный анализ).

В общем виде схему применения этих методов можно проиллюстрировать на одном из наиболее распространенных методов такого типа — на факторном анализе. Основная гипотеза факторного анализа заключается в том, что каждый объект-стимул можно описать как линейную комбинацию небольшого числа основных факторов. Число и характер этих факторов определяют *априорно* выделенные “существенные” параметры объектов. На основе измерений выделенных характеристик объектов строятся корреляционные или ковариационные матрицы, анализ которых приводит к локализации стимулов в пространстве основных факторов, которые интерпретируются как субъективные шкалы. Каждую субъективную шкалу соотносят с физическими параметрами стимула, выявляя связь по типу одномерной психофизической функции. В случае обнаружения однозначной связи между субъективными измерениями и физической переменной, задачу можно считать решенной. Любой новый стимул будет расположен на субъективной шкале просто путем вычислений по результатам физических измерений.

Но перенос этого принципа на субъективное шкалирование сложных стимулов порождает новую проблему. Даже в случае, когда для сложного стимула известны физические параметры, с которыми связано изменение субъективных оценок, совсем не так просто установить однозначную связь между стимулом и реакцией. Например, восприятие аппертурных цветов традиционно определяется такими субъективными характеристиками, как цветовой тон, насыщенность и яркость. Известны и определяющие их физические параметры: длина волны светового излучения, спектральный состав излучения (чистота) и его интенсивность. Казалось бы, достаточно иметь три одномерные функции типа $R=f(S)$, чтобы описать ощущение такого сложного стимула, как цвет. Но все оказывается значительно сложнее, поскольку изменение длины волны излучения приводит к изменению не только цветового тона, но одновременно меняется и другая субъективная характеристика — насыщенность, например. Или изменение интенсивности светового излучения приводит к изменению яркости, но вместе с тем меняется и ощущение цветового тона, хотя длина волны излучения не меняется, и т.д. Таким образом, простой набор одномерных психофизических функций не гарантирует описания субъективного изменения многомерного стимула.

§6. Многомерное шкалирование

Приведенный пример с цветовыми стимулами, когда непосредственные физические измерения не объясняют однозначно субъективных шкал, является наиболее типичным случаем (и наиболее интересным); именно в этом случае и применяются разрабатываемые в последние годы новые методы измерения, называемые *многомерным шкалированием* (Торгерсон, 1958; Шепард, 1962; Крускэл, 1964).

Ясно, что когда люди оценивают сложное качественное свойство объектов, такое как эмоциональное выражение лица, или когда они оценивают общее сходство сложных объектов, они ведут себя так, как если бы мерили объекты сразу по нескольким субъективным шкалам, а не по одной. Комбинируя определенным образом субъективные меры, они и осуществляют сложное суждение, подобное оценке психологического качества. Иначе говоря, сложную субъективную шкалу можно представить как систему нескольких простых субъективных шкал. Тот факт, что люди используют для объяснения некоторого качества зачастую более чем одну физическую шкалу, наводит на мысль, что субъективные шкалы также могут быть составными. Можно представить себе далее, что свои измерения по субъективным шкалам люди осуществляют какими-то не всегда осознанными методами комбинирования характеристик объектов. Поэтому вполне вероятно, что некоторые из этих субъективных шкал не соответствуют однозначно простым физическим шкалам. Применение методов одномерного шкалирования, описанных выше, оставляет мало надежды получить полезную информацию об этих “составных” шкалах.

В конечном итоге становится очевидным, что модель, которая походит на многомерные модели в том, что может трактоваться как система шкал, и которая не требует, чтобы эти шкалы определялись *заранее* до анализа данных (как, например, в случае с факторным анализом), будет обладать большой ценностью. Подобная модель может применяться непосредственно к оценочным данным, чтобы получить основные субъективные шкалы безотносительно к любым предположениям о физических коррелятах этих шкал.

Для решения таких проблем и развивались методы многомерного шкалирования. При использовании этих методов предполагается, что в основе сложных суждений человека лежит система из нескольких субъективных шкал, которая и формирует субъективное пространство. Когда испытуемых просят сравнить пару объектов, они ориентируются на различия между объектами по каждой субъективной шкале, и итоговая оценка различия есть величина, производная от различий по каждой шкале. В качестве модели системы субъективных шкал обычно используется геометрическое пространство, точки которого представляют исходные стимулы. Оси геометрического пространства интерпретируются как субъективные шкалы, а шкальные величины каждого стимула — как значения координат соответствующей точки. Предполагается, что если стимулы представить как точки пространства, то субъективные оценки различий определенным образом соотносимы с расстояниями между точками в субъективном пространстве. Конкретный вид связи между субъективными различиями и межточечными расстояниями в каждом случае может быть различным. Он определяется используемой моделью субъективного расстояния, но в данном случае понятно, что два стимула, сильно различающиеся между собой, будут расположены на далеком расстоянии друг от друга в пространстве, а сходные стимулы расположатся рядом.

В примере, который приводился, субъектам предлагали оценивать пары цветовых карт по заданному качеству (красоте). В случае многомерного шкалирования от субъекта требуется аналогичная оценка, но не самого отдельного признака, а оценки степени общего *сходства или различия* между парами стимулов, по которым строится шкала межстимульных различий. Результаты подобных измерений содержат в себе всю информацию о структуре множества стимульных точек в субъективном пространстве. Вопрос заключается в том, как ее оттуда извлечь?

Рассмотрим гипотетический пример многомерного шкалирования цветоразличения. Предположим, что мы предъявили испытуемому три цветные карты (*A, B и C*). Карта *A* окрашена в белый цвет, карта *C* — в желтый и карта *B* — в оранжевый.

Оценки различий, которые произвел испытуемый для всех карт, следующие:

Стимулы	***A***	***B***	***C***
A	**0**	**5**	**4**
B	**5**	**0**	**3**
C	**4**	**3**	**0**

Возьмем в качестве модели субъективного расстояния *евклидову метрику* и допустим, что оценки различий связаны с расстояниями в евклидовом пространстве прямой пропорциональностью. Если данные оценки действительно основаны на одномерной шкале, мы должны расположить точки, представляющие наши цветные карты, вдоль одной оси так, чтобы расстояние между точками соответствовало оценкам различий. Это значит, что, памятуя о предположении, что оценки различий прямо соотносимы с расстояниями в субъективном пространстве, расстояние между какими-то двумя точками должно равняться сумме расстояний от этих точек до третьей. Но для данных вышеприведенной таблицы это следствие никак не может быть выполнено. Из этого следует, что три данных цвета нельзя расположить на одномерной шкале. Поскольку стимулов всего три, то они располагаются, как минимум, в двухмерном пространстве (рис.1), иначе говоря, оцениваются по двум шкалам. Определяя минимальную размерность пространства, которая снимает несовместимость полученных оценок, метод многомерного шкалирования позволяет обнаружить число необходимых субъективных шкал, лежащих в основе сложных суждений.

Полученные две оси субъективного пространства легко интерпретировать как две субъективные характеристики цветовых стимулов: шкалу цветового тона и шкалу цветовой насыщенности. Эта интерпретация наглядно следует из структуры цветовых точек в пространстве (рис. 1), построенном по данным этой таблицы. В соответствии с такой интерпретацией можно выбрать систему ортогональных координат и вычислить проекции точек на оси. Эти значения будут прямо соответствовать шкальным значениям цветового тона и насыщенности анализируемых стимулов.

Конечно, в подобном случае перед исследователями остается проблема физической интерпретации этих субъективных шкал, и хотя многомерное шкалирование не предлагает однозначного решения этой проблемы, но оно все-таки является более полезным, чем процедура одномерного шкалирования. Прежде всего, определяя действительные субъективные шкалы, которые субъекты используют для оценивания объектов, а не априорный набор шкал, метод многомерного шкалирования дает исследователю больше шансов выявить физические свойства объектов, которые человек реально использует как основу для субъективных измерений. Другими словами, отношение между каждой субъективной и физической шкалой может быть определено раздельно. Кроме этого, явно определяя субъективные шкалы (даже если соответствующие физические шкалы неизвестны), эти процедуры дают полезную информацию о том, какие физические измерения необходимо сделать, чтобы попытаться найти физический коррелят субъективной шкалы. И последнее, вся эта информация получается без обращения к традиционным методам шкалирования — более сложным и трудоемким.

Важное значение по отношению к задачам прикладного характера имеют особенности многомерного шкалирования, связанные с выявлением не только структуры субъективного пространства стимулов, но и возможностей определения

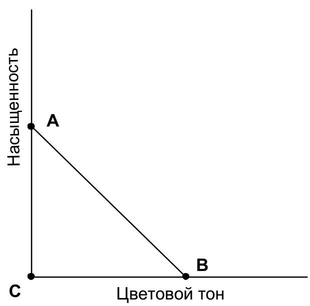


Рис.1 Гипотетический пример расположения трех цветов (A, B и C) в двухмерном евклидовом пространстве

тонких индивидуальных различий между самими испытуемыми (Клифф, 1973; Терехина, 1975; Виш и Кэрролл, 1974).

Рассмотрим это на примере тех же цветовых карт. Допустим, что оценивая различия между картами, одни испытуемые будут больше ориентироваться на различия по цветовому тону (рис. 2), другие — по насыщенности (рис. 3), а третьи — одинаково на те и на другие (рис.1).

Если все три группы испытуемых расположить в двумерном пространстве цветовой тон — насыщенность, мы получим следующую картину (рис. 4).

Испытуемые группы 3 придают большое значение различиям в цветовой насыщенности, испытуемые группы 1 — раз-

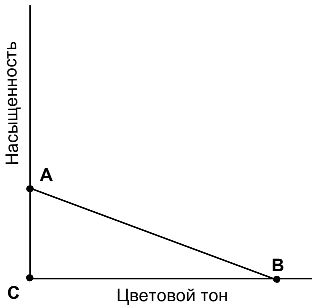


Рис. 2. Гипотетический пример субъективного пространства восприятия трех цветов (A, B и C) испытуемыми, больше ориентирующимися при различении цветов на цветовой тон

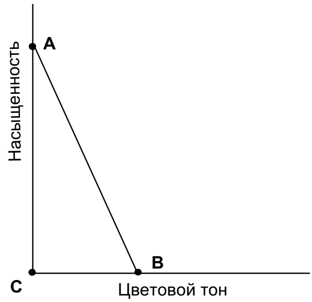


Рис. 3. Гипотетический пример субъективного пространства восприятия трех цветов (A, B и C) испытуемыми, больше ориентирующимися при различении цветов на насыщенность

личиям в цветовом тоне и игнорируют различия в насыщенности. Испытуемые 2 группы занимают среднее положение. Легко представить, что между этими группами возможны промежуточные варианты, и в таком случае можно использовать, например, угол наклона луча, проходящего из точки пересечения осей через данный класс испытуемых, как некоторую основу для классификации испытуемых (рис. 4). Измеряя этот угол относительно оси цветового тона, мы сможем разделить группы испытуемых по их предрасположению к цветовому тону или насыщенности.

Естественно, что этот пример, как и все предыдущие, в значительной степени упрощен и схематизирован. Это сделано для того, чтобы подчеркнуть суть данной методики в применении к анализу индивидуальных различий, когда модель многомерного шкалирования предназначена не только для определения субъективных признаков, лежащих в основе суждений, но и для выделения более значащего из этих признаков, вносящего наибольший вклад в индивидуальные суждения.

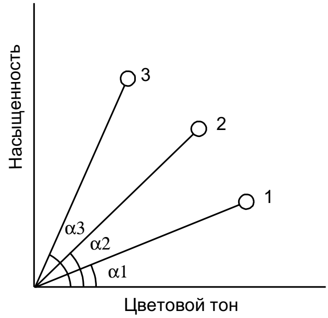


Рис. 4 Пространство индивидуальных различий трех групп испытуемых

Изложение сложных вопросов в простой и краткой форме неизбежно связано с разрывами в логической структуре. Этот недостаток будет компенсирован в следующих разделах при более детальном и строгом изложении моделей и процедур шкалирования. Однако более общий взгляд на исследуемый предмет обладает преимуществом охвата одновременно всех главных частей и связей, как это делает крупномасштабная карта. Чтобы еще раз подчеркнуть взаимосвязь различных методов шкалирования, подытожим в виде отдельных пунктов главные моменты излагаемого предмета:

1. Психофизические измерения начинались с построения одномерных субъективных шкал (процедуры субъективного шкалирования Фехнера и Стивенса). Необходимым условием для осуществления субъективных измерений было наличие соответствующей физической шкалы. Физические измерения служили основанием и критерием истинности для субъективных измерений. В то же время применение этих моделей было чрезвычайно ограничено этим условием.

2. Принципиально новым шагом в психофизике явилась разработка методов субъективных измерений, не требующих предварительного построения физической шкалы стимула (модель шкалирования Терстоуна). Эти методы значительно расширили сферу применения субъективных измерений, включив в нее шкалы, не имеющие явных физических коррелятов. Существенным ограничением этих методов яв-

ляется обязательная одномерность измеряемой субъективной характеристики стимула.

3. Последующее развитие методов шкалирования связано с построением сложных многомерных шкал. Но практический прогресс соединялся в данном случае с методическим отступлением. В основание многомерных субъективных шкал вновь кладутся определенные характеристики объекта, которые должны быть заданы априорно, т.е. еще перед началом исследования (модели факторного анализа и др.).

4. Наконец, свою наиболее развитую форму психофизические измерения получили в моделях многомерного шкалирования, когда производимый анализ многомерных реакций на сложные стимулы не связан с предварительным физическим анализом стимуляции, а ориентирован исключительно на внутреннюю структуру суждений.

ЧАСТЬ I
ЛОКАЛИЗАЦИЯ ТОЧКИ НА ШКАЛЕ (НОЛЬМЕРНОЕ ШКАЛИРОВАНИЕ)

Глава 1. МЕТОДЫ ИЗМЕРЕНИЯ ПОРОГОВ

Исторически сложилось так, что первыми методами психологических измерений были методы, позволяющие определять локализацию точки на психологической шкале. Их появлением мы обязаны Г.Т. Фехнеру, пытавшемуся с их помощью разрешить *психофизическую проблему* — выяснить закон соответствия психического образа и вызвавшего его физического воздействия. Согласно Фехнеру, через абсолютный порог задается начальная точка отсчета на психологической шкале, а через разностный порог вводится единица измерения на ней.

Под порогом всегда подразумевается некое критическое значение, разделяющее исследуемый ряд явлений на 2 класса. *Абсолютный порог* — то минимальное значение в континууме стимулов, выше которого раздражитель всегда воспринимается. *Разностный порог* — то минимальное различие в выраженности определенного физического параметра стимулов, превышение которого приводит к восприятию их различия.

Для решения основной задачи Фехнером была предпринята разработка методов измерения порога. Предложенные им три метода измерения порога добрую сотню лет были единственными методами определения чувствительности и до сих пор признаются классическими. Знакомство с ними и в наши дни составляет основу подготовки психолога-экспериментатора. И дело не только и даже не столько в том, что порог как мера чувствительности — психической способности воспринимать, чувствовать, реагировать — широко используется в различных областях психологических исследований. Главное в другом. Хотя пороговые, как и любые другие психофизические измерения, если их рассматривать в содержательном аспекте, представляют собой частный случай психометрических измерений, в пороговых методах — наиболее простых измерительных процедурах — уже нашли отражение основные трудности количественного определения переменных и способы их преодоления: вариабельность измеряемых величин и использование средних значений для их характеристики; их вероятностный характер, влияние многочисленных и далеко не всегда контролируемых экспериментатором факторов и введение уравновешивающих их действие процедур. Оба указанные качества — простота и характерность — определяют место пороговых методов в курсе обучения методам психологических измерений.

Развитие методов в науке обычно связано с появлением новых проблем. Главной и единственной задачей классической психофизики было изучение закона соответствия между психическими и физическими переменными. Основное внимание уделялось стимульным переменным, поскольку молчаливо предполагалось *отсутствие влияния* на ответы испытуемого в эксперименте таких *несенсорных факторов* как изменение мотивации, получение дополнительной информации об экспериментальной ситуации и т.д. Эти предположения реализовались в процедурных особенностях пороговых измерений и представлениях о пороге как мере чувствительности. Характерной чертой трех классических пороговых методов является довольно большое разнообразие стимулов, применяемых в эксперименте в качестве *независимой переменной*, и отсутствие какого бы то ни было контроля упомянутых выше несенсорных факторов, фактически всегда включенных в эксперимент. Те статистические показатели, которые приняты в этих методах в качестве пороговых мер, на самом деле являются, строго говоря, *мерами исполнения*, т.к. определяются не только уровнем чувствительности испытуемого, но и теми несенсорными факторами, которые управляют выбором его ответа. Несмотря на это, такие качества пороговых методов, как простота, меньшие затраты времени на измерение, а также удобство выражения пороговых мер в физических единицах обеспечивают этим методам широкое применение в современной исследовательской практике.

Современная психофизика (в отличие от классической), решая проблему измерения чувствительности, основное внимание уделяет *процессу решения* снсорной задачи (выбора испытуемым ответа) в типичной для порогового эксперимента ситуации отсутствия ясных отчетливых впечатлений от действия стимула. Это и определило характерные черты нового класса методов, детальная разработка которых осуществлена в последнее тридцатилетие — *методов обнаружимости сигнала*. Общим для всех методов этого класса является резкое обеднение стимульной ситуации (сведение ее всего до двух стимулов) и варьирование в качестве независимой переменной факторов, управляющих выбором ответа испытуемого. Описанию этих методов посвящена глава 2 данного раздела.

§1. Метод минимальных изменений

Данный метод является единственным среди методов измерения чувствительности, который дает знание величины порога *в ходе самого измерения*. В процедуре этого метода прямо отразилось понимание порога как барьера, разделяющего стимульный ряд на два класса ощущаемых и неощущаемых стимулов или их разностей.

1. Измерение абсолютного порога (RL) методом минимальных изменений.

Процедура. Существует несколько вариантов процедуры измерения этим методом. Рассмотрим процедуру Вундта. Каждая проба начинается сигналом “Внимание”, после которого с постоянным интервалом (0,5 - 1,5 секунды) предъявляется стимул, например, пятно света при определении *абсолютной световой чувствительности* в полной темноте. Как правило, испытуемому разрешается только *две категории ответов* (“Да”, “Нет”; “Вижу”, “Не вижу” и т.п.), форма которых точно оговаривается в инструкции испытуемому. Испытуемый отвечает, его ответ регистрируется. Предъявление стимулов осуществляется *нисходящими и восходящими рядами*. В первом случае степень выраженности определенного параметра стимула, чувствительность к которому измеряется, постепенно уменьшается от максимума до минимума, во втором — наоборот. Обычно измерение абсолютного порога начинается с нисходящего ряда стимулов, т.е. с отчетливо воспринимаемого стимула, изменяемый параметр которого с каждым шагом последовательно уменьшается. За порог в этом ряду принимается значение стимула, находящегося *в середине межстимульного интервала* между тем стимулом, который еще воспринимается, и тем, который впервые не воспринимается, т.е. середина того интервала, в котором произошла *первая смена категории ответа* испытуемого. В нисходящем ряду определяется порог исчезновения ощущения — L_l,в восходящем — порог появления — L_h (L — от латинского limen — порог). Чаще всего они не совпадают вследствие существования *систематической ошибки*.

Систематические ошибки бывают двух типов. Это так называемая *ошибка привыкания*, когда испытуемый продолжает повторять тот же ответ, что и на предыдущем шаге, хотя порог уже пройден и стимул в нисходящем ряду уже не вызывает ощущения, и *ошибка ожидания или предвосхищения* — ошибка противоположного толка. Для того, чтобы *сбалансировать* любую из этих ошибок, если они появляются, применяется: 1) уравновешивание числа тех и других рядов путем их чередования — нисходящие и восходящие ряды предъявляются парами, 2) требование от испытуемого ответа на каждый шаг изменения стимула в ряду. Для контроля за тщательностью работы испытуемого используется еще один экспериментальный прием — *изменение длины стимульных рядов* от пары к паре за счет смещения в случайном порядке начального и конечного значения стимулов в ряду. Эта предосторожность служит для предупреждения возможности повторения испытуемым своих ответных реакций на основе простого отсчета от начала и конца ряда определенного количества шагов изменения стимула[1].

При выборе *величины шага изменения* стимула надо учитывать следующие моменты. При уменьшении величины шага падает дисперсия ответов (Геррак, 1970), а, следовательно, и порогов в восходящих и нисходящих рядах, что позволяет сократить число пар рядов, не изменяя заданной точности измерения порога. Однако, уменьшение величины шага приводит к увеличению количества шагов в каждом отдельном ряду, т.е. к удлинению ряда и, следовательно, опыта в целом. Оптимальный размер шага является результатом компромисса между стремлением к большой точности в оценке порога и нежеланием делать опыт очень длинным и утомительным.

Необходимое число измерений (пар рядов) определяется требуемой *точностью измерения* и степенью разброса получаемых в эксперименте данных. Его можно вычислить по формуле:

$$n = \frac{U_p^{2}\sigma^{2}}{\delta^{2}}, \qquad (1)$$

где n — число измерений; U_p — квантиль нормального распределения, соответствующий заданной доверительной вероятности в определении порога; σ — дисперсия пороговых значений; δ — требуемая точность в определении порога.

Поскольку до начала опытов дисперсия пороговых значений неизвестна, для определения требуемого числа измерений необходимо провести предварительные пробные измерения, чтобы “прикинуть” величину дисперсии.

<u>Обработка результатов.</u> За абсолютный порог принимается среднее арифметическое всех найденных в течении

---

[1] Действительно, ведь сообразив в первых 2—3 рядах, что пороговая величина стимула соответствует где-то четвертому шагу изменения стимула в восходящем ряду из 10 шагов, испытуемый может в остальной части опыта заниматься только счетом: четвертый шаг снизу — “порог”, шестой сверху — тоже “порог”, и т.д.

опыта порогов появления и исчезновения и рассчитывается как :

$$RL = \frac{1}{N}\sum_{i=1}^{N} L_i \quad , \qquad (2)$$

где RL — средний абсолютный порог (обозначение RL — аббревиатура от немецкого “Reiz Limen”); L_i — значение единичного порога в каждом стимульном ряду, как в восходящем, так и в нисходящем; N — общее число рядов.

Вариативность работы испытуемого оценивается средним квадратическим (стандартным) отклонением σ_L:

$$\sigma_L = \sqrt{\frac{\sum_{i=1}^{N}(L_i - RL)^2}{N-1}} \quad , \qquad (3)$$

Статистическая ошибка, которая допускается, при вычислении в опыте абсолютного порога, оценивается *стандартной ошибкой среднего значения:*

$$m_{RL} = \frac{\sigma_L}{\sqrt{N-1}} \qquad (4)$$

2. Измерение дифференциального порога (DL) методом минимальных изменений.

<u>Процедура.</u> В этом случае все особенности метода и процедура остаются почти теми же, что и при определении абсолютного порога. Единственное изменение процедуры состоит в том, что одновременно с переменным стимулом испытуемому предъявляется эталон или стандартный стимул — S_{st}, который задает тот уровень исходного раздражителя, относительно которого выясняется величина разностного порога. В силу того, что ощущения различия стимулов у испытуемого могут быть различны, естественно разрешить испытуемому давать *три*

$S_{пер.}$	↓	↑	↓	↑	↓	↑	↓	↑
15	+				+			
14	+		+		+		+	
13	+	+	+	+	+	+	+	
12	+	+	+	=	+	+	+	+
11	+	=	=	=	+	=	=	+
10	=	=	=	=	=	=	=	=
9	=	=	=	=	=	=	=	=
8	=	=	=	=	=	-	=	=
7	=	=	=	=	=	-	=	=
6	=	=	=	-	-	-	=	=
5	=	-	=	-	-	-	=	-
4	=	-	-	-	-		-	-
3	-		-		-		-	-
2	-						-	
1							-	

Рис.1. Фрагмент протокола опыта по измерению дифференциального порога:
<+>, <=> и <-> — ответы испытуемого. Точками отмечены пороги в восходящих (↑) и нисходящих (↓) рядах

категории ответов, а именно “больше”, “меньше”, “равно”. Ответ “не знаю”, “сомневаюсь” обычно отождествляется с ответом “равно”. За порог принимается значение стимула, соответствующее середине межстимульного интервала, где впервые произошла смена категории ответа: от “больше” к “равно” и от “равно” к “меньше” в нисходящем ряду, а в восходящем ряду от ответа “меньше” к ответу “равно” и от ответа “равно” к ответу “больше”. Таким образом, при измерении разностного порога определяются *четыре* значения порога (по два в каждом ряду). Это верхний порог — L_h в восходящем и нисходящем рядах ($L_{h\uparrow}$ и $L_{h\downarrow}$) и нижний порог — L_l в восходящем и нисходящем рядах ($L_{l\uparrow}$ и $L_{l\downarrow}$). Таким образом, в каждом ряду мы находим две пороговые точки: *верхний и нижний разностные пороги*. На рис.1 они помечены точками в каждом ряду. Этот рисунок иллюстрирует правило установления пороговой точки в нисходящих и восходящих рядах ответов испытуемого.

Обработка данных. Сначала находим значения верхнего разностного порога путем усреднения всех верхних порогов в каком бы ряду они не стояли:

$$L_h = \frac{\sum_{i=1}^{n} (L_{h\uparrow} + L_{h\downarrow})}{n} \quad , \qquad (5)$$

где L_{h-} и $L_{h\text{Ї}}$ — значения верхних порогов в восходящем и нисходящем рядах, а n — число пар рядов.

Аналогичным образом вычисляем нижний разностный порог:

$$L_l = \frac{\sum_{i=1}^{n} (L_{l\uparrow} + L_{l\downarrow})}{n} \quad . \qquad (6)$$

Верхний и нижний пороги ограничивают *интервал неопределенности* — IU (от английского “Interval of Uncertainty”), т.е. ту зону стимульного ряда, где преобладают ответы равенства. Иначе говоря, интервал неопределенности — это та зона стимулов, которая сверху ограничена стимулом, в среднем едва заметно отличающимся от эталонного, как больший, а снизу — стимулом, в среднем едва заметно отличающимся от эталонного, как меньший. Понятно поэтому, что IU содержит две различительные ступени или два едва заметных различия, т.е. *равен двум дифференциальным порогам* DL (от немецкого “Differenz Limen”):

$$IU = L_h - L_l \quad , \qquad (7)$$

$$DL = \frac{IU}{2} = \frac{L_h - L_l}{2} \quad . \qquad (8)$$

Стимул, находящийся в средней точке интервала неопределенности, всегда оценивается как равный эталону, т.е. является субъективным эквивалентом эталона и потому по-

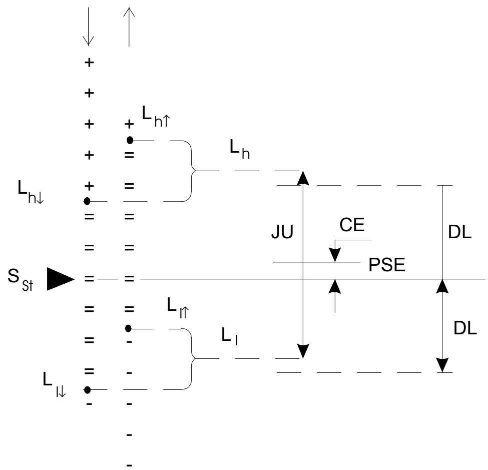


Рис.2. Соотношение основных пороговых показателей, оцениваемых в ситуации измерения дифференциального порога в методе минимальных изменений

лучил название *точки субъективного равенства* PSE (от английского "Point of Subject Equality"):

$$PSE = \frac{L_h + L_l}{2} \quad . \qquad (9)$$

IU, как правило, несимметричен, поэтому довольно часто PSE не совпадает со значением эталона. Степень несовпадения эталона PSE характеризуется так называемой *константной ошибкой,* CE (от английского "Constant Error"), которая определяется следующим равенством:

$$CE = PSE - S_{st} \qquad (10)$$

Если константная ошибка больше нуля, то эталон переоценивается, если она меньше нуля, то эталон недооценивается. Таким образом, CE характеризует величину и направление смещения зоны субъективного равенства относительно объективного равенства. Соотношение этих основных психофизических понятий, которые используются и в других методах, иллюстрируется схемой, приведенной на рис.2.

3. Варианты метода минимальных изменений.

Объединение пары рядов в один ряд. В этом случае восходящий и нисходящий ряды предъявляются без перерыва. Достоинство этого варианта в том, что он обеспечивает некоторое сокращение времени. Его существенным недостатком является увеличение при такой системе подачи стимулов нерегулярности ответов, обусловленной тем, что второй ряд в паре начинается со стимула, вызывающего слабое, неуверенное ощущение различия.

Процедура “вверх-вниз” (метод лестницы). Этот вариант метода границ, предложенный *Корнсвитом (1962)*, предполагает использование двух вариантов ответов. Суть его состоит в том, что как только происходит смена категории ответа, допустим, смена ответа “слышу” на ответ “не слышу”, так сразу же происходит смена направления изменения стимула, т.е. переход от нисходящего ряда к восходящему до следующей смены категории ответа. Этот вариант метода относится к так называемым *адаптивным методам* пороговых измерений, и, как правило, реализуется на компьютере, который отслеживает ответы испытуемого и соответствующим образом регулирует изменение стимуляции. В этих методах процедура тестирования строится таким образом, что предъявление стимулов подстраивается (“адаптируется”) под ответы испытуемого, и изменение стимуляции происходит в достаточно узком околопороговом диапазоне (рис. 3).

Достоинством этой процедуры является экономичность, вместе с тем она имеет ряд недостатков. Один из них состоит в том, что эта модификация метода применима только к измерению абсолютного порога. Дифференциальный порог может измеряться этим методом только в разных двух сериях, а это плохо из-за временных колебаний чувствительности. Второй недостаток состоит в том, что испытуемый быстро замечает порядок чередования ощущаемых и неощущаемых стимулов, что вызывает эф-

Интенсивность стимула																					
10	+																				
9		+																			
8			+																		
7				+																	
6					+					+											
5						+										+		+			
4							+		-		+		+		-		-				+
3												-							+		
2								-						+						-	
1																					

Рис.3. Запись ответов испытуемого при изменении интенсивности стимуляции в опыте по измерению абсолютного порога методом “лестница”:
при ответах <+> (“да”) интенсивность стимула уменьшается, при ответах <—> (“нет”) — увеличивается

фект ожидания, распространяющийся по горизонтали, т.е. переносится с одного стимула на другой. Гилфорд (1954) отмечает, что этот эффект является таким сильным источником стабилизации результатов (несенсорным источником!), равного которому нет ни в каком другом методе. Величину этой стабилизации трудно измерить и каким-либо образом скорректировать. Поэтому эта процедура применима только в случаях, когда исследователь может удовлетвориться очень грубым, но зато быстрым определением порога. На практике этот метод часто применяется для скрининговых исследований слуховой чувствительности.

Метод едва заметного различия (ЕЗР). Эту модификацию иногда полностью отождествляют с методом границ. Это не совсем точно. Суть метода ЕЗР сводится к следующему: испытуемому вместе с эталонным стимулом предъявляют ряд переменных. Задача испытуемого состоит в том, чтобы указать то значение стимула, которое едва заметно отличается от эталона. Существенное отличие метода ЕЗР от метода границ состоит в том, что в методе границ испытуемый определяет два порога — порог появления и исчезновения ощущения различия — ЕЗР и ЕНЗР, т.е. едва незаметного различия. В методе ЕЗР определяется только одна точка — всегда ЕЗР. Эти методы тождественны только тогда, когда изменение стимуляции начинается от равенства переменного и эталонного стимулов. В том случае, когда изменение стимуляции начинается от заметного неравенства к равенству, испытуемый определяет точку исчезновения ощущения различия (ЕНЗР), опираясь на сенсорный эталон, хранящийся в памяти.

§2. Метод средней ошибки

В отечественной литературе этот метод известен также под названием *метод воспроизведения, метод подгонки, метод подравнивания и метод установки.* Этот метод отличается от других пороговых методов двумя процедурными особенностями — испытуемый *сам регулирует* величину изменяемого параметра стимула; стимул может принимать любое значение в заданном диапазоне, т.е. его изменения *непрерывны.* Фехнером этот метод предназначался для измерения дифференциальной чувствительности. Позднее он стал использоваться для измерения абсолютной чувствительности, хотя, по мнению Фехнера, метод средней ошибки (МСО) не позволяет прямо измерить порог; он дает меру, пропорциональную чувствительности. Вместе с тем это единственный метод, в котором субъективный эквивалент эталона определяется *непосредственно в процедуре измерения.* Второй отличительной особенностью этого метода является наиболее естественная для испытуемого процедура определения равного эталону стимула путем *собственноручного подравнивания.* Благодаря этим свойствам МСО довольно часто применяется в исследованиях восприятия. Именно этот случай применения метода средней ошибки стал хрестоматийным (Гилфорд, 1954; Вудвортс и Шлосберг, 1971).

1. Применение метода средней ошибки для измерения дифференциального порога.

Процедура. При измерении дифференциальной чувствительности испытуемому предъявляются одновременно два стимула, эталон — S_{st} и переменный S_{var}, величину которого может изменять испытуемый. Аппаратура должна позволять *плавную регулировку* изменяемого параметра переменного стимула. Задача испытуемого состоит в *подравнивании* переменного стимула к эталону. Испытуемому дается установка на точность, а не на быстроту воспроизведения эталона. Никаких ограничений на свободу движений при регулировке стимула в процессе подравнивания не вводится. Подравнивание должно начинаться то от большего, чем эталон, значения, то от меньшего. Чтобы исключить для испытуемого возможность осуществлять подравнивание на основе одного только *кинестезического впечатления*, необходимо в обоих случаях *менять начальные точки*. Обычно бывает достаточно выбрать три заметно различающиеся начальные значения переменного стимула, большие и меньшие, чем эталон, и чередовать их, применяя в течение опыта равное число раз. В силу наличия в *протетических континуумах* (Стивенс, 1960)[1] пространственной ошибки, в опыте должно быть сделано равное число проб с положением эталона слева и справа от переменного стимула или сверху—снизу от него.

Обработка данных. Для качественного анализа результатов опыта полезно построить гистограмму распределения подравниваний, что несложно сделать на компьютере с помощью практически любого статистического пакета. Кроме того целесообразно построить график распределения результатов подравниваний во времени. Наглядное представление результатов опыта в графической форме несомненно поможет глубже и содержательнее проанализировать не только

---

[1] Этот термин введен Стивенсом для обозначения такого континуума стимулов, для которого количественные изменения стимула вызывают количественные изменения соответствующих ощущений; например: интенсивность звука и громкость.

различные стратегии решения испытуемым сенсорной задачи, но и наглядно оценить динамику его работы.

В качестве статистических мер, необходимых для оценки пороговых показателей, в МСО принято характеризовать полученное распределение чаще всего средним арифметическим (см. формулу (2) в предыдущем параграфе) и реже — медианой. В качестве мер разброса используются стандартное отклонение (см. формулу (3) в предыдущем параграфе) и реже — полумежквартальный размах. Очень редко в настоящее время используется такая мера изменчивости полученных данных, как среднее отклонение или средняя ошибка:

$$F = \frac{\sum_{i=1}^{N} \left| x_i - M \right|}{N} \quad , \qquad (11)$$

где x_i — одно из значений в ряду подравниваний; M — среднее арифметическое подравниваний; n — количество подравниваний.

Меры чувствительности, используемые в МСО. В литературе можно найти разноречивые рекомендации в отношении мер чувствительности, которыми следует пользоваться в пороговых измерениях с помощью МСО. В результате экспериментов по подравниванию исследователь получает распределение установок испытуемого, которое характеризуется локализацией на стимульной оси и отмеченными выше показателями изменчивости. По мнению автора МСО Фехнера, при измерении этим методом исследователь получает не прямую оценку порога, а только пропорциональную ей величину, коей является один из показателей разброса — *средняя ошибка*. Логическим основанием для этого могло служить соображение о том, что в соответствии с инструкцией (подравнять переменный стимул к эталону) центр распределения подравниваний должен характеризовать субъективный эквивалент эталона. По смыслу введенных выше понятий он является точкой субъективного равенства (PSE). Вместе с тем, чем более размыт, расплывчат субъективный эквивалент эталонного стимула, чем меньше испытуемый может отличить его от соседних значений, тем ниже чувствительность. По-видимому, Фехнер придавал именно такой психофизический смысл этому показателю разброса и поэтому описанный им метод был назван методом средней ошибки. Однако в целом ряде исследований были получены разные типы локализации распределения подравниваний на стимульной оси — смещенное и несмещенное относительно положения эталона. В связи с этим ряд исследователей (Челпанов, 1925; Осгуд, 1954; Торгерсон, 1958; Вудвортс и Шлосберг, 1965; Бардин, 1976) предлагают использовать в качестве меры чувствительности также и величину отстояния субъективного эквивалента эталона (центра распределения подравниваний) от эталона. Обосновывается это предложение тем, что чем ниже чувствительность испытуемого, тем более далекие стимулы он принимает равными эталону, поэтому эти два разные показателя как бы характеризуют чувствительность с разных сторон, и потому оба имеют право на существование. Вместе с тем никто из этих авторов не обращает внимания на то обстоятельство, что по смыслу введенных выше определений предлагаемая ими мера оценки чувствительности как разность значений точки субъективного равенства и эталона является константной ошибкой: $CE = PSE - S_{st}$.

Экспериментальными исследованиями показано, что константная ошибка определяется главным образом систематическими ошибками измерения, такими как *пространственные* и *временные* ошибки[1].

---

[1] Пространственная ошибка связана с различным расположением в пространстве эталонного и переменного стимулов; например, эталон может переоцениваться или недооцениваться в зависимости от того, где расположен переменный стимул — снизу или сверху от него. Временные ошибки обусловлены порядком предъявления в паре эталонного и переменного стимулов; например, если эталон предъявляется первым, то он может переоцениваться.

Цель измерения и выбор адекватной инструкции для испытуемого. Причиной получения разных типов локализации распределения подравниваний на стимульной оси является то, что классический вариант инструкции “подравнять переменный стимул к эталону” дает испытуемому большую свободу в ее трактовке, поскольку в переходной зоне от значений стимула меньших, чем эталон, до значений больших, чем эталон, существует целый ряд стимулов, кажущихся равными эталону — *интервал неопределенности* (IU), а инструкция не уточняет, какую именно точку в этом ряду должен искать испытуемый. Экспериментальные исследования последних лет (Михалевская, Скотникова, 1978) позволили дать обоснованную интерпретацию психофизического смысла статистических показателей, получаемых в методе средней ошибки, и показали, что при определенных модификациях инструкции метод средней ошибки позволяет обоснованно и точно определить все основные психофизические показатели, а именно, интервал неопределенности, точку субъективного равенства и дифференциальный порог. Оказалось, что психофизический смысл среднего значения подравниваний определяется тем, какую сенсорную задачу решает испытуемый, т.е. тем, какая *инструкция или самоинструкция* им принята. Для измерения границ интервала неопределенности, и, следовательно, разностного порога как половины интервала неопределенности, испытуемому должно быть указано на поиск точки *первого равенства* переменного стимула и эталона. В таком эксперименте, где подравнивание начинается от стимулов, заметно больших и заметно меньших, чем эталон, результаты подравнивания представляют собой *бимодальное* (двугорбое) распределение (рис. 4).

Если отдельно обработать данные, полученные в пробах, где исходные значения переменного стимула были заметно меньше и где они были заметно больше, чем эталон, то центры этих распределений (их средние арифметические) будут характеризовать *нижнюю и верхнюю границы интервала неопределенности*. Следовательно, при такой организации процедуры МСО становится возможным получить

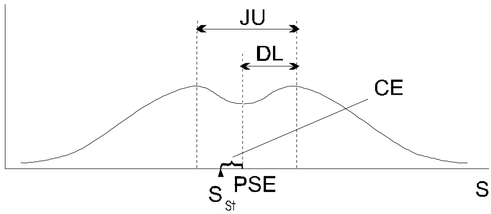


Рис. 4. Распределение результатов подравниваний испытуемым переменного стимула к эталонному при инструкции искать точку первого равенства и чередовании исходных значений изменения переменного стимула от заметно больших и заметно меньших, чем эталон:
по оси абсцисс – величина стимула, по оси ординат – частота подравнивания стимула к стандартному

оценку дифференциального порога, т.е. снимается то ограничение этого метода, которое имел в виду Фехнер.

Если исследователя интересует локализация субъективного эквивалента эталона, т.е. точка субъективного равенства, то испытуемый должен подравнивать *к центру* зоны неразличения (равенства) переменного стимула и эталона. Экспериментально доказано, что среднее распределения подравниваний, полученного в результате выполнения испытуемым такой инструкции, локализуется в центре интервала неопределенности и совпадает с точкой субъективного равенства.

Величина другого показателя метода средней ошибки — стандартного отклонения подравниваний — зависит преимущественно от сенсорной способности и характера двигательных действий испытуемого по подравниванию. *Стандартное отклонение* (σ) является индивидуально устойчивой характеристикой испытуемого в *метатетических континуумах* стимулов (Стивенс, 1960)[1] и не зависит от локализации среднего подравниваний в зоне перехода от различения к нераз-

[1] В метатетическом континууме количественные изменения стимулов вызывают качественные изменения ощущений, например: длина волны светового стимула и цветовой тон.

личению. Поэтому при условии обучения испытуемых определенным оптимальным двигательным действиям по подравниванию стандартное отклонение, хотя и является мерой исполнения (т.е. зависит не только от чувствительности сенсорной системы, но и от особенностей процесса принятия испытуемым решения, включенного в этот эксперимент), может служить хорошей операциональной оценкой дифференциальной чувствительности к метатетической стимуляции.

2. Применение метода средней ошибки для измерения абсолютной чувствительности.

В этом случае испытуемый регулирует величину стимула, первоначально вызвавшего отчетливое ощущение, до тех пор, пока не установит такое его значение, при котором он впервые утрачивает ощущение воздействия стимула. Если установка начинается с явно неощущаемой величины стимула, то испытуемый должен найти такое его значение, при котором ощущение впервые появляется. Обычно рекомендуется для оценки абсолютного порога использовать такие *меры центральной тенденции,* как медиана и среднее. Меры изменчивости (межквартильный размах и стандартное отклонение) в данном случае характеризуют только вариативность установок. В случае получения бимодального распределения за оценку абсолютного порога следует брать середину расстояния между двумя экстремумами аналогично определению точки субъективного равенства в задаче измерения разностного порога.

3. Общая оценка и область применения метода средней ошибки.

Является общепризнанным, что метод средней ошибки дает *наиболее низкие* значения порога по сравнению с другими методами. Это объясняется, по-видимому активной сенсомоторной деятельностью субъекта, т.е. возможностью регулировки самим испытуемым стимуляции и связанным с этим привлечением других источников информации (кинестезии) для решения стоящей перед ним задачи, а также большим, как правило, временем действия стимула, а, следовательно, возможностью более полного извлечения информации из стимуляции.

Процедура подравнивания очень естественна и легко принимается всеми испытуемыми — взрослыми и детьми. Это расширяет область ее применения по сравнению с другими методами. Метод подравнивания оказывается незаменим при оценке чувствительности во всех случаях, когда сенсорная чувствительность оператора является средством (орудием), используемым оператором при решении профессиональных задач в процессе трудовой деятельности, как, например, у фотометриста, определяющего плотность вещества путем подгонки к эталону, или токаря, обтачивающего деталь с точностью до микрона.

Наиболее адекватно применение МСО в тех случаях, когда требуется оценка точки субъективного равенства. Именно этим объясняется довольно широкое применение метода в шкалирующих процедурах.

Существенно ограничивает область применения метода средней ошибки необходимость обеспечения плавной регулировки стимуляции, что, в свою очередь, может быть достаточно сложной технической проблемой.

§3. Метод постоянных раздражителей

Другие названия этого метода — *метод констант, частотный метод, метод истинных и ложных случаев*. Метод состоит в предъявлении испытуемому ряда стимулов, неизменных в течение всего опыта, и название отсюда — *метод постоянных раздражителей (МПР)*, метод констант. В случае измерения разностного порога предъявляется стандартный стимул и сравниваемый с ним. В силу того, что параметры стандартного и сравниваемого стимулов в течение всего опыта неизменны, каждый из сравниваемых стимулов образует со стандартным постоянную разницу. Отсюда еще одно название этого метода — метод постоянных разниц. Непосредственным результатом опыта являются *частоты ответов*, из которых значения порога находятся *вычислительным путем*. Эта особенность определила еще одно название этого метода — метод частот.

Метод констант пользуется репутацией самого точного и надежного, поскольку сама процедура метода предусматривает такую организацию стимуляции, которая *исключает ошибки привыкания и ожидания*. Возможность накопления большой статистики ответов, связанная с ограничением числа постоянных раздражителей, применяемых в измерении, *повышает надежность измерения* порога этим методом. Универсальность метода констант обусловлена, по-видимому, двумя обстоятельствами. Во-первых, он ставит менее жесткие требования к выходным устройствам задающей аппаратуры, чем метод средней ошибки, поскольку высокоточную дискретную регулировку выходного сигнала получить технически существенно проще. Это значительно расширяет область применения МПР. Во-вторых, дискретность стимуляции позволяет использовать, кроме суждений, и другие ответные реакции организма, например, вегетативные, электроэнцефалографические, сосудистые и др. Эти реакции отличаются двумя важными для измерения чувствительности свойствами: 1) не поддаются произвольному контролю (без специальной тренировки), 2) их величина изменяется градуально. Использование этих реакций существенно расширяет область приложения МПР, поскольку обеспечивает его применение в тех случаях, когда исследователю невозможно (или неудобно) использовать речевой ответ для измерения порога (например, в случаях симуляции, у детей, еще не овладевших речью, у животных). Кроме того, применение непроизвольных реакций позволяет увеличить объем информации, извлекаемой из опыта, поскольку информация об изучаемом процессе содержится не только в факте появления/не появления реакции, но и в ее величине, форме и скрытом периоде, поэтому возрастает количество сведений, которое может быть извлечено из каждой градуальной реакции.

Давая общую характеристику метода констант, нельзя не отметить еще одного момента. Метод констант занимает особое место среди классических методов измерения чувствительности в связи с тем, что почти все теоретические построения психофизики относительно пороговой проблемы для своего экспериментального подтверждения обращались к этому методу. Он оказался наиболее гибким, получаемые этим методом результаты находили объяснение в русле самых различных психофизических концепций.

1. Определение разностного порога методом констант.

Процедура. В предварительных испытаниях экспериментатор ориентировочно определяет *пороговую зону*, т.е. тот диапазон различия стимулов, на границах которого испытуемый начинает практически всегда ощущать отличие эталонного стимула от сравниваемого. Затем экспериментатор выбирает в пределах этой зоны ограниченный ряд стимулов, которые будут сравниваться с эталоном (чаще всего 5—7). Выбор производится с таким расчетом, чтобы самый слабый среди них вызывал у испытуемого ответ "больше" в 5—10% случаев, а самый сильный — в 90—95%. Сравниваемые стимулы выбираются так, чтобы расстояния между ними на стимульной оси были одинаковыми. Последнее требование обеспечивает некоторое упрощение статистической обработки данных и является просто требованием удобства. При определении разностного порога стимулы предъявляются парами — эталон и сравниваемый — одновременно или последовательно. Стимульная последовательность, составленная из пар стимулов, является по своим свойствам *случайной, но сбалансированной*: каждая пара предъявляется равное число раз, частота предъявления каждой пары распределена на последовательности равномерно. Естественно, что эта последовательность составляется до опыта и испытуемому неизвестна. Обычно в опыте каждая пара стимулов повторяется 20—200 раз.

В экспериментальной практике используются два разных способа объединения стимулов в пары: 1) место эталона в паре меняется по случайному закону; 2) место эталона и сравниваемого стимула в паре фиксированы. Первый вариант решения имеет то преимущество, что позволяет компенсировать постоянные ошибки типа пространственной и временной в ходе самого эксперимента. Сильным аргументом в пользу второго способа является уменьшение вариативности результатов опыта за счет уменьшения колебаний критерия при выборе испытуемым ответа в каждой отдельной пробе. По-видимому, следует предпочитать второй способ, а пространственную ошибку можно учесть, если в одной стимульной последовательности эталон предъявляется слева, а в другой — справа. Аналогичным образом можно выявить и временную ошибку.

В каждой пробе, т.е. при предъявлении пары стимулов, испытуемый должен вынести суждение, возникло ли ощущение различия и каково оно. В методе констант используются две (“больше”, “меньше”) или три категории ответов (“больше”, “меньше”, “равно”). В любом случае порог вычисляется из пропорций суждений разного рода на каждую пару стимулов.

Психометрическая функция.

Рассмотрим случай, когда испытуемый дает две категории ответов — “больше” и “меньше”. Обозначим S_{st} — стандартный стимул, а S_{var} — сравниваемый по исследуемому параметру (один из постоянных стимулов). Если S_{var} существенно меньше S_{st}, то испытуемый почти никогда не дает ответ “больше”, если же S_{var} значительно больше S_{st}, то почти всегда испытуемый дает ответ “больше”. В промежутке между этими двумя значениями при увеличении изменяемого параметра стимула пропорция ответов “больше” плавно возрастает от 0 до 1. Поэтому пропорцию ответов “больше” удобно использовать при представлении результатов эксперимента в виде графика, называемого *психометрической функцией.*

Если в эксперименте предъявить достаточно большое число раз несколько пар S_{var}, S_{st} и представить полученные данные на графике, где по абсциссе отложена физическая мера стимулов, а по ординате для каждого стимула указана пропорция ответов “больше”, то точки, описывающие ответные данные, образуют кривую, имеющую, как правило, *S-образную* форму. Если выбрать некоторое новое значение сравниваемого стимула, которое лежит между уже опробованными, и повторить эксперимент, то соответствующая ему новая точка придется между двумя старыми. Это дает основание заключить, что для любой пары стимулов S и S_{st} существует вероятность $P(S_{var})$ ответа “S_{var} больше S_{st}”. *Психометрической функцией* называется такая функция P аргумента S, которая является монотонной, дифференцируемой и ограничена нулем и единицей (Урбан, 1907). Оценкой ее значений служат пропорции ответов “больше”. Из дифференцируемости и ограниченности нулем и единицей можно сделать вывод о существовании соответствующей ей дифференциальной функции распределения. Принятые в психофизике изображения психометрической и дифференциальной кривой распределения, полученных в эксперименте, проведенном методом констант с двумя категориями ответов, представлены на рис. 5.

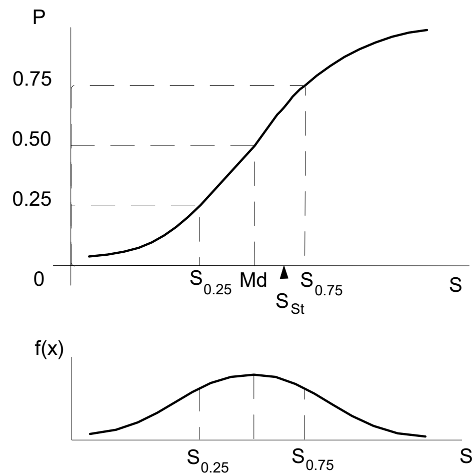


Рис. 5. Психометрическая функция и соответствующая ей дифференциальная кривая:

оси абсцисс на обоих графиках — интенсивность сравниваемого стимула; ось ординат на верхнем графике — вероятность ответов “больше”, ось ординат на нижнем графике — плотность вероятности ответов “больше”

Форма психометрической кривой. S-образная форма психометрической кривой допускается как пороговыми теориями Фехнера и Блеквелла, так и теориями непрерывности, хотя интерпретация ее в том и в другом случае различна. Основная суть любой пороговой теории сводится к утверждению о существовании порога как реального принципа работы сенсорной системы. Порог понимается буквально как барьер, критическое значение в континууме раздражений. Если бы значение порога было стабильно во времени, то психометрическая кривая имела бы вид линейной ступенчатообразной функции. Этого никогда не бывает. Ее S-образная форма объясняется тем, что порог флуктуирует во времени случайным образом. Различные варианты альтернативных теорий (Дельбеф, 1883; Мюллер, 1896; Ястров, 1888), отвергающие существование порога, исходили из предположения, что ощущение является непрерывной функцией, зависящей от двух переменных — интенсивности раздражителя и степени предрасположенности человека к его восприятию. Поскольку последняя зависит от случайного сочетания множества трудно учитываемых факторов, то их баланс является случайной величиной и имеет нормальное распределение. Именно поэтому и психометрическая кривая имеет S-образный вид *интегральной функции нормального распределения*. Фехнер (1860) также считал, что психометрическая функция является интегральной функцией нормального распределения; эта точка зрения получила название *фи-гамма гипотезы*. В старых работах классической психофизики φ (фи) использовалась для обозначения стимулов, а γ (гамма) — для обозначения ответов. Терстон (1928) полагал, что поскольку согласно закону Вебера различительная ступень растет с увеличением стимула, психометрическая кривая приобретает положительную асимметрию, пропорциональную дроби Вебера. Психометрическая кривая нормализуется, если взять логарифмический масштаб по стимульной оси (φlogγ-гипотеза). Различие психометрических кривых, полученных в пороговых экспериментах, столь незначительно, что трудно отдать предпочтение одной из этих гипотез.

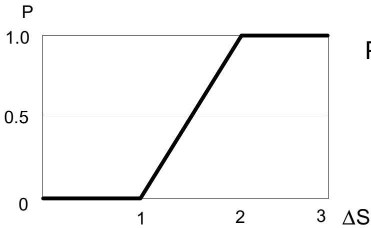


Рис. 6. Психометрическая функция, предсказанная нейроквантовой теорией Стивенса

Нейроквантовая теория Стивенса, являющаяся по существу пороговой, предсказывает прямолинейную психометрическую кривую, представленную на рис.6.

Согласно этой теории, изменение в ощущении замечается всегда, когда дополнительное возбуждение, вызванное приращением стимула, увеличивается на величину, равную одному *нервному кванту*. Порог различения отстоит от стандартного стимула согласно этой теории на 1,5 стимульных интервала, соответствующие нервному кванту. Фактор случайности в этой теории воплощается не в колебаниях порога, а в случайной величине остаточного возбуждения, суммируясь с которой добавочное возбуждение, вызванное приращением стимула, приводит к генерации нервного кванта. Предполагается, что условная единица стимульной оси служит физическим аналогом величины кванта. Прямолинейность психометрической кривой обусловлена равномерным распределением величины остаточного возбуждения.

Параметры психометрической кривой. Как и в других пороговых методах для характеристики распределения результатов измерения в МПР используются меры центральной тенденции (медиана — Md и среднее арифметическое — M) и меры изменчивости (полумежквартильный размах — Q и стандартное отклонение — σ). Перпендикуляр из медианы дифференциальной кривой распределения делит площадь под кривой пополам. Поскольку площадь под кривой равна единице, медиане соответствует стимул, для которого вероятность ответа “больше” равна 0,5:

$$Md = S_{0.5} \quad . \qquad (15)$$

Полумежквартильный размах определяется как полуразность Q_3 и Q_1[1]:

$$Q = \frac{S_{0.75} - S_{0.25}}{2} \quad . \qquad (16)$$

В последние годы в практике психофизических исследований стали часто использоваться среднее арифметическое распределения — М и стандартное отклонение — σ_s. Как известно, в симметричных распределениях Md и М совпадают, а меры изменчивости строго соотнесены:

$$\boldsymbol{\sigma}_s = 1.483Q \quad . \qquad (17)$$

<u>Психофизический смысл параметров психометрической кривой.</u>

Интервал неопределенности оценивается через межквартильный размах (Q_3 - Q_1):

$$IU = S_{0.75} - S_{0.25} \ . \qquad (18)$$

Точка субъективного равенства определяется как медиана: PSE=Md. *Константная ошибка* имеет место в случае несовпадения медианы со стандартом и равна:

$$CE = Md - S_{st} \quad . \qquad (19)$$

Разностный порог определяется в эксперименте с двумя категориями ответов как половина интервала неопределенности и соответствует полумежквартильному размаху психометрической кривой, построенной по ответам "больше" или "меньше". Обозначим его Q(2), где цифра в скобках указывает на количество категорий ответа, а индекс Q подчеркивает, что порог характеризуется мерой разброса:

$$DL(2) = Q(2) = \frac{S_{0.75} - S_{0.25}}{2} \quad . \quad (20)$$

---

[1] Напомним, что Q_1, Q_2, Q_3 и Q_4 находятся на оси абсцисс психометрической функции в точках, соответствующих вероятностям P(>) = 0.25, 0.5, 0.75 и 1.0.

Психофизические показатели в эксперименте с тремя категориями ответов.

При использовании трех категорий ответов испытуемого в методе констант — "больше", "меньше" и "равно" — психометрические кривые ответов "больше" и "меньше" не являются зеркальными и потому должны рассматриваться обе. Результаты 3-категориального эксперимента представлены на рис.7 и 8.

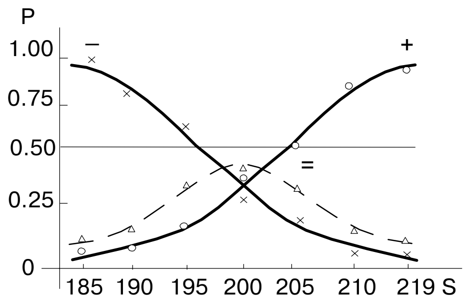


Рис.7. Психометрические кривые, полученные в эксперименте с тремя категориями ответов:
<+> — ответы "больше", <-> — "меньше", <=> — "равно"

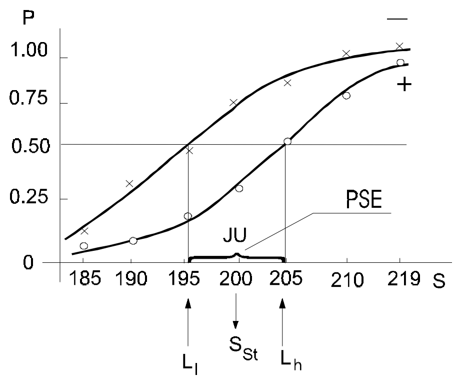


Рис. 8. Те же психометрические кривые, что и на рис. 7:

кривая ответов "меньше" получена путем дополнения вероятностей ответов "меньше" до 1, т.е. проведена через точки, полученные как $(1\text{-}P_{(-)})$; PSE — точка субъективного равенства; IU — интервал неопределенности; L_l и L_h — нижний и верхний пороги различения, соответственно

В соответствии с принятым операциональным определением порога как 50% точки (см. пункт 2 в этом параграфе), которое можно полностью применить к трехкатегориальному эксперименту, медиана психометрической кривой ответов “меньше” является оценкой нижнего разностного порога, а медиана ответов “больше” — верхнего разностного порога; расстояние между ними характеризует интервал неопределенности (IU), центр которого является точкой субъективного равенства (PSE). За величину разностного порога одни исследователи (Урбан, 1907; Бардин, 1976) предлагают считать согласно принятому в методе границ определению половину интервала неопределенности, т.е.

$$DL(3) = \frac{L_h - L_l}{2} \quad , \qquad (21)$$

где DL(3) — обозначение указанной оценки разностного порога; L_h и L_l — величины верхнего и нижнего разностного порога, соответственно.

Другие исследователи предлагают использовать в качестве пороговой меры различения полумежквартильный размах психометрической кривой ответов “больше” или “меньше”. Обозначим эту оценку порога Q(3):

$$Q(3) = \frac{S_{0.75} - S_{0.25}}{2} \quad . \qquad (22)$$

Эта оценка разностного порога, по мнению Каллера (1928), не зависит от частоты появлений ответов “равно” и появилась как следствие неудовлетворенности психофизиков первой мерой DL(3), поскольку величина интервала неопределенности сильно зависит от стремления испытуемого употреблять нейтральные ответы. В самом деле, при увеличении частоты ответов “равно” увеличивается величина интервала неопределенности (см. рис. 9), а следовательно, и разностного порога, если его оценивать как DL(3). Фернбергер (1931) показал, что величина IU сильно зависит от инструкции, с помощью которой можно управлять частотой ответа “равно”. Он заключил, что испытуемые различаются по частоте употребления ответов “равно” частично в силу различия темперамента, частично — в результате различия в инструкциях. Следовательно, величина IU характеризует вклад скорее процесса решения, чем собственно сенсорной чувствительности.

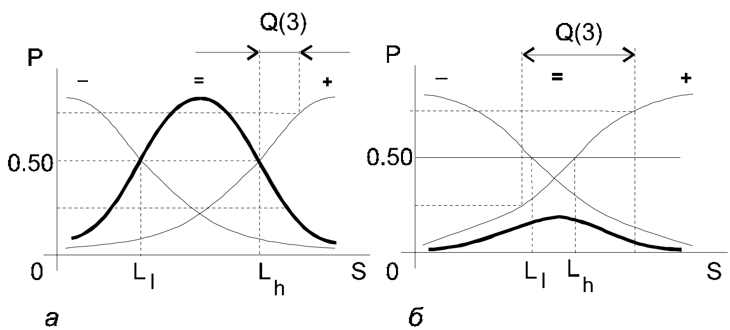


Рис. 9. Психометрические функции, полученные в МПР с тремя категориями ответов:
а — большое количество ответов “равно”, *б* — незначительное количество ответов “равно”

Однако и вторая оценка порога различения Q(3) столь же подвержена критике. Гилфорд (1954), сравнивая оценки порогов по DL(3) и Q(3) показал, что они измеряют разные величины. Если испытуемый уменьшает число ответов “равно”, крутизна психометрических кривых ответов “больше” и “меньше” уменьшается, т.е. одновременно с уменьшением IU и DL(3) увеличивается Q(3) (см. рис.9*б*). Если же испытуемый по каким-либо причинам увеличивает число нейтральных ответов, соотношение величин DL(3) и Q(3) изменяется в противоположном направлении (см. рис.9*а*). Отсюда Гилфорд делает обоснованный вывод, что две оценки, которые меняются в противоположных направлениях, не могут служить мерой одного и того же.

Результатом описанной дискуссии явился отказ психофизиков от использования трех категорий ответов при измерении порогов методом констант: испытуемому либо вообще не разрешается использовать нейтральные ответы в ходе опыта, либо в случае разрешения нейтральные ответы делятся между ответами “больше” и “меньше”. Вопрос о том, как делить ответы — поровну или пропорционально количеству ответов двух других категорий — много обсуждался, но исследователи так и не пришли к согласованному мнению.

При использовании двухкатегориальной системы ответов единственной используемой оценкой порога реакции является величина Q(2), поэтому везде, где это только возможно, при работе МПР применяются только две категории ответов.

Рекомендация отказаться от трехкатегориальной системы ответов при измерении чувствительности методом констант не всегда приемлема. В тех случаях, когда требуется оценка различия сложных многомерных стимулов, испытуемый затрудняется классифицировать свои ощущения в терминах “больше” — “меньше”, поскольку при изменении одного параметра стимула может меняться сразу несколько сенсорных признаков воздействия, и испытуемый, “соскальзывая” с одного признака на другой, может испортить эксперимент. По-видимому, наиболее подходящим для оценки восприятия сложных стимулов, наиболее часто встречающихся в прикладных исследованиях, является метод, позволяющий испытуемому выносить суждение о различии стимулов, не “привязываясь” к какому-либо одному признаку, и обеспечивающий такую организацию эксперимента, которая позволила бы уменьшить загрубляющее оценку собственно чувствительности влияние несенсорных факторов. Модификация метода АБХ, предложенная Индлиным (1978), по-видимому, удовлетворяет этим требованиям (см. пункт 3).

2. Определение абсолютного порога методом констант.

Процедура измерения абсолютного порога от измерения разностного порога методом констант отличается только тем, что в каждой пробе испытуемому предъявляется один из нескольких (обычно 5—9) постоянных стимулов, на который испытуемый дает один из двух возможных ответов. Определение стимульного диапазона, количества предъявляемых стимулов, величины межстимульного интервала осуществляется исходя из тех же соображений, которые учитывались при организации измерения дифференциального порога. Порядок предъявления стимулов также строится как сбалансированно случайный.

По полученным в эксперименте частотам ответов на каждый из постоянных стимулов строится психометрическая кривая. За абсолютный порог принимается так называемая 50-процентная точка кривой, т.е. мера центральной тенденции (среднее M или медиана Md, чаще медиана). Почему 50-процентная точка берется в качестве пороговой меры? С точки зрения пороговой концепции эта точка есть медиана распределения моментальных значений порога, т.е. значений абсолютного порога в те моменты времени, когда происходит измерение. С точки зрения классической теории непрерывности ответ испытуемого есть функции двух переменных — величины стимула (чем больше, например, интенсивность стимула, тем чаще ответ “Да”) и баланса благоприятных и неблагоприятных факторов разной природы. 50-процентной точке соответствует минимальное значение стимула, вызывающего ощущение только при балансе благоприятных и неблагоприятных факторов.

Меры изменчивости, описывающие полученное распределение, полумежквартильный размах — Q и стандартное отклонение — σ, характеризуют надежность оценки порога.

Естественно, измеряя абсолютный порог, мы должны отдавать себе отчет в том, что это не столько порог “чистого” ощущения, сколько порог реакции, т.е. величина, на которую влияют и несенсорные факторы. В частности, истинное значение порога ощущения может искажаться за счет влияния случайного угадывания. Для корректировки таких ответов в рамках пороговой концепции Блэквеллом (1953) была предложена так называемая “поправка на случайный успех”. Согласно Блэквеллу, вероятность правильного ответа “Да” складывается из вероятности истинного восприятия предъявляемого стимула (P_c) и вероятности случайного угадывания неощущаемого воздействия. Последняя величина равна вероятности ответа “Да” ($P_{\text{"yes"}}$) при отсутствии стимула (иначе называемая ложной тревогой — P_{fa}), умноженной на вероятность отсутствия ощущения при воздействии стимула, т.е.

$$P_{\text{"yes"}} = P_c + P_{fa}(1 - P_c), \tag{23}$$

откуда истинная вероятность правильных ответов определяется из результатов эксперимента следующим образом:

$$P_c = \frac{(P_{\text{"yes"}} - P_{fa})}{1 - P_{fa}} . \tag{24}$$

Следует помнить, что исходной посылкой Блэквелла было отрицание какой-либо сенсорной основы ответов угадывания, с чем трудно согласиться, поскольку известно, что далеко не весь опыт рефлексируется человеком.

Для того, чтобы воспользоваться “поправкой на случайный успех”, необходимо ввести в эксперимент так называемые пустые пробы (пробы-ловушки), когда после сигнала “Внимание” экспериментатор не предъявляет стимула. Возникающие в этих пробах ответы “Да” позволят оценить вероятность ложных тревог.

Рассмотрим пример определения абсолютного порога методом констант. Измеряется *пространственный порог тактильного восприятия* — то минимальное расстояние между двумя раздражаемыми точками кожи, при котором испытуемый в 50% случаев дает ответ “два” и в 50% — ответ “один”. Выбрав участок кожи, на котором будет определяться порог, экспериментатор делает несколько предварительных замеров *эстезиометром*, используя, например, процедуру метода границ, для того, чтобы грубо определить пороговую зону, внутри которой некоторые предъявления стимула вызывают ответ “Два”, а некоторые другие предъявления стимула — ответ “Один”. Экспериментатор выбирает 5 стимулов таким образом, что наименьший стимул вызывает ответ “Два” приблизительно в 5% случаев, а наибольший — в 95%. Интервалы между стимулами равны. Предъявляются стимулы в сбалансированно-случайном порядке. Каждый стимул предъявляется 100 раз. Экспериментальные данные приведены в таблице 1. По этим данным строится психометрическая кривая. Для этого на графике по абсциссе откладывается физический параметр стимула — расстояние между раздражаемыми точками в мм, а по ординате — пропорции ответов. Психометрическая кривая нашего примера приведена на рис.10.

Таблица 1

Результаты эксперимента по определению пространственного порога тактильного восприятия

Чрезвычайно редко случается так, что одному из сти-

Расстояние между стимулами S, мм	8	9	10	11	12
Пропорция ответов “Два” ($P_{\text{”два”}}$)	0.01	0.05	0.29	0.66	0.93
Результат преобразования $P_{\text{”два”}}$ в $Z_{\text{”два”}}$	-2.33	-1.55	-0.55	0.41	1.48

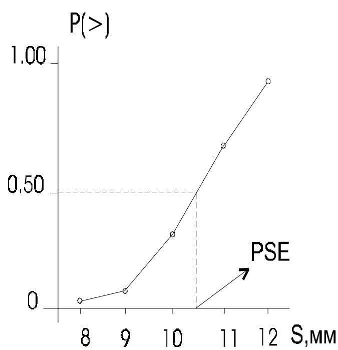


Рис. 10. Психометрическая кривая, построенная по результатам эксперимента по определению пространственного порога тактильного восприятия:

точками показаны экспериментальные результаты. Полученная кривая является хорошим приближением к интегральной кривой нормального распределения (по Гилфорду, 1954) мулов соответствует пороговая пропорция ответов: $P_{\text{"два"}} = 0{,}5$. Чаще всего соответствующую порогу точку приходится определять по полученной психометрической кривой. Графическим или вычислительным путем можно найти значения медианы (и среднего), характеризующих величину абсолютного порога (в нашем примере RL=10,57 мм) и меры вариативности — квартили Q_3, Q_1 и стандартного отклонения — σ.

Очевидно, что точность оценки порога обусловлена прежде всего "хорошестью" *аппроксимации* экспериментально полученных точек гладкой кривой. К сожалению, математически корректное решение задачи *подгонки* точки непросто. Поэтому на практике используются два варианта построения психометрической функции: 1) с помощью *линейной интерполяции* отдельных участков психометрической функции в линейных координатах; либо 2) вся психометрическая функция аппроксимируется функцией нормального распределения, которое в нормальных координатах является *прямой* линией. Рассмотрим оба эти случая обработки экспериментальных данных.

Обработка экспериментальных данных в методе констант

Способ линейной интерполяции. Этот способ не обеспечивает высокую точность, но зато крайне прост. *Линейная интерполяция*[1] основывается на представлении психометрической функции в виде отрезков прямой, которые проводятся между полученными точками. Этот случай представлен на рис. 11.

[1] Метод линейной интерполяции основан на допущении, что на участке между двумя экспериментальными точками психометрическая функция может быть приблизительно представлена в виде прямой. Такое предположение в известной степени правомерно, поскольку на интересующем нас участке между Q_1 и Q_3 психометрическая функция действительно похожа на прямую линию.

Простейшим и наиболее часто используемым является *графический способ* нахождения значений медианы и квартилей. Если на графике провести горизонтальные линии на уровне пропорций ответов, равных 0.5, 0.25, 0.75, то их пересечения с построенной психометрической кривой дадут, сотвeтственно, значения Md, Q_1 и Q_3, которые считываются с оси абсцисс в физических величинах стимула. Естественно, при использовании графического способа обработки результатов следует построить психометрическую функцию на координатной бумаге, выбрав достаточно крупный масштаб.

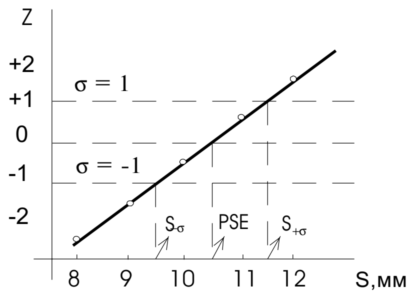


Рис. 11. Психометрическая функция, построенная по экспериментальным точкам с использованием метода линейной интерполяции

Те же значения могут быть получены и *расчетным путем* по следующим формулам (фактически эти формулы вытекают из решения прямоугольных треугольников):

Медиана психометрической кривой определяется как

$$Md = S_l + \frac{(S_h - S_l)+(0.5 - P_l)}{P_h - P_l} \quad , \qquad (25)$$

где S_l — величина ближайшего к 50-процентной точке стимула, лежащего ниже ее, S_h — величина стимула, лежащего непосредственно выше 50-процентной точки, P_l и P_h — соответствующие указанным выше стимулам пропорции ответов.

Первый и третий квартили вычисляются по формулам:

$$Q_1 = S_{l1} + \frac{(S_{l1} - S_{h1})(0.25 - P_{l1})}{P_{h1} - P_{l1}}, \qquad (26)$$

где S_{l1} — величина стимула, лежащего непосредственно ниже 25-процентной точки,

S_{h1} — величина стимула, лежащего непосредственно выше 25-процентной точки,

P_{l1} и P_{h1} — соответствующие указанным выше стимулам пропорции ответов.

$$Q_3 = S_{l3} + \frac{(S_{l3} - S_{h3})(0.75 - P_{l3})}{P_{h3} - P_{l3}}, \qquad (27)$$

где S_{l3} — величина стимула, лежащего непосредственно ниже 25-процентной точки; S_{h3} _ величина стимула, лежащего непосредственно выше 25-процентной точки; P_{l3} и P_{h3} — соответствующие указанным выше стимулам пропорции ответов.

В нашем примере Md = 10,57 мм, Q_1 = 9,83 мм, Q_3 = = 11,33 мм.

Недостатками способа линейной интерполяции являются:

1) расточительность, так как из всех полученных в эксперименте данных используется только часть — например, для определения Md достаточно иметь две точки;

2) отсутствие возможности получить *точную* оценку показателей разброса – дисперсии или межквартильного размаха – Q. Если в эксперименте используется больше двух стимулов, можно определить Q_1 и Q_3, а если допустить, что распределение частот ответов является нормальным, то можно найти и величину стандартного отклонения через соот-

ношение σ = 1,483Q. Однако при широком диапазоне используемых стимулов и относительно малом их числе (около 5, как в нашем примере) оценка Q будет не очень точной, следовательно, и значение σ также.

Способ нормальной интерполяции. Если сделать более строгое допущение о форме психометрической функции, а именно, что она является функцией нормального распределения, и если выразить масштаб оси ординат в единицах стандартного отклонения этого распределения, то психометрическая функция, имеющая S-образную форму в линейных координатах, *превращается в прямую линию*. После этого появляется возможность найти все интересующие исследователя параметры прямой, аналогично тому, как это делалось в случае линейной интерполяции. Но для этого нужно прежде всего преобразовать пропорции ответов P с помощью таблиц нормального распределения в значения Z, представляющие собой нормированные по стандартному отклонению расстояния от стимульных точек до медианы. После *преобразования P в Z* экспериментальные точки на графике, где по абсциссе отложен физический параметр стимула S, а по ординате — Z, могут быть аппроксимированы[1] прямой линией, которая проводится "на глазок" (этот способ хотя и прост, но чаще всего дает лишь грубое приближение), либо рассчитывается с помощью *метода наименьших квадратов*. Этот метод позволяет получить не только наилучшую аппроксимацию, но и статистически строго оценить степень "хорошести" подгонки полученной прямой к экспериментальным точкам.

Определение медианы, представленной в z-координатах психометрической функции, возможно графичес-

[1] Термин "аппроксимация" экспериментальных точек какой-либо функцией означает процедуру представления (моделирования) набора эмпирических точек в виде определенной математической функции. В данном случае предполагается, что, если психометрическая функция — это функция нормального распределения, то в нормальных координатах она будет иметь вид линейной функции. Очевидно, что "хорошеть" арппоксимации экспериментальных точек линейной функцией будет одновременно служить показателем адекватности принятого предположения о нормальности распределения.

ким и расчетным путем. За абсолютный порог (и PSE при измерении двухкатегориальным методом констант разностного порога) принимается величина стимула, которой соответствует Z = 0. Стандартное отклонение определяется как такая величина стимула, для которой Z = +1 или Z = -1[1]. Через стандартное отклонение можно найти и величину полумежквартильного размаха — Q, т.к. их связь в случае нормального распределения описывается равенством

$$Q=0{,}674\sigma . \qquad (28)$$

Для иллюстрации этого способа обработки обратимся к нашему примеру (см. табл. 2). Графическое представление зависимости величины $Z_{\text{„два"}}$ от физического параметра стимула (т.е. психометрическая функция в нормальных координатах) приведено на рис. 11.

Определение с помощью графиков параметров психометрической функции способом нормальной интерполяции не требует преобразования в z-координаты, если имеется в наличии вероятностная бумага. Способ изготовления такой бумаги подробно описан (Бардин, 1976).

Все необходимые пороговые показатели могут быть определены и аналитическим путем с помощью соответствующих формул. Для этого можно воспользоваться двумя методами.

Во-первых, можно применить уже известный нам метод линейной интерполяции (теперь в нормальных координатах), который фактически является аналогом простого графического решения, когда мы не производим строгого построения аппроксимирующей прямой. Расчет параметров психометрической прямой производится по формулам (29), (30) и (31):

---

[1] Фактически шкала z-оценок и является шкалой единиц стандартного нормального отклонения — σ . Точка z=0 соответствует нулевому отклонению от среднего (медианы), точки z=1 или -1 — отклонению от среднего на 1σ вправо или влево, соответственно.

$$RL = Md = \frac{z_h \cdot S_l - z_l \cdot S_h}{z_h - z_l} \quad , \qquad (29)$$

где z_l и z_h — самые близкие к нулю отрицательная и положительная величины z, соответственно; S_l и S_h — стимулы, соответствующие z_l и z_h (т.е. величины ближайшего подпорогового и надпорогового стимулов).

Для оценки величины стандартного отклонения следует взять разность между точками на стимульной оси, соответствующими z=1 или z=-1 и величиной порога — RL. Эти точки можно вычислить так:

$$S_{\sigma+} = \frac{S_{h+}(1 - z_{l+}) - S_{l+}(1 - z_{h+})}{z_{h+} - z_{l+}} \quad , \qquad (30)$$

где z_{l+} и z_{h+} — ближайшие значения z, меньшие и большие +1, соответственно; S_{h+} и S_{l+} — стимулы, соответствующие z_{l+} и z_{h+} (т.е. ближайшие значения стимулов, ниже и выше $S_{\sigma+}$).

$$S_{\sigma-} = \frac{S_{l-}(1 + z_{h-}) - S_{h-}(1 + z_{l-})}{z_{h-} - z_{l-}} \quad , \qquad (31)$$

где z_{l-} и z_{h-} — ближайшие значения z, меньшие и большие -1, соответственно; S_{h-} и S_{l-} — стимулы, соответствующие z_{l-} и z_{h-} (т.е. ближайшие значения стимулов, ниже и выше $S_{\sigma-}$).

Оба значения $S_{\sigma+}$ и $S_{\sigma-}$ вычисляются в связи с тем, что полученная в эксперименте психометрическая кривая далеко не всегда является очень хорошим приближением к кривой нормального распределения, и эти значения могут расходиться. Поэтому обычно для оценки разброса используется их среднее. В нашем примере вычисления по приведенным формулам дали следующие величины:

RL = 10.57 мм, $S_{\sigma+}$ и $S_{\sigma-}$ = 0.98 мм.

Во-вторых, воспользовавшись методом наименьших квадратов, можно построить *наилучшую прямую*, проходящую через экспериментальные точки. Эта задача решается просто в любом статистическом пакете путем выполнения процедуры построения простой линейной регрессии. Вычислив таким образом коэффициенты *a* и *b* линейной функции $y=ax+b$, мы без труда найдем неизвестные "x" по известным "y" (z=0 , z=1 или z=-1). Понятно, что поскольку точки $S_{\sigma+}$ и $S_{\sigma-}$ будут симметричны относительно RL, то достаточно вычислить лишь одну из них.

3. Варианты метода констант

Метод приращения. Особенностью экспериментальной процедуры является непрерывное предъявление испытуемому стандартного стимула, к которому периодически добавляются приращения. Испытуемый отвечает, заметил ли он приращение, в терминах, например, "Да"-"Нет". Разностным порогом является приращение стимула, заметное в 50% случаев. В методе приращений измеряется разностный порог реакции, представляющий собой половину интервала неопределенности. Сомнения в отношении возможности использования интервала неопределенности в качестве показателя различения уже высказывались выше.

В экспериментах, проводимых в поддержку нейроквантовой теории, практикуется вариант метода приращений, при котором в каждой экспериментальной серии используется лишь одна величина приращения. Наличие перерывов между экспериментальными сериями с разными величинами приращений является недостатком этого метода, поскольку допускает направленное изменение характеристик испытуемого в отношении приращений разной величины.

Метод АБХ. В этом методе испытуемому предъявляются последовательно три стимула: первый обозначается А, второй — Б, третий — Х. Первые два стимула различаются величиной исследуемого параметра; в качестве третьего стимула (Х) используется либо А, либо Б. Испытуемый должен ответить, какой из стимулов был X. Метод АБХ при условии запрещения нейтральных ответов сводится к двухкатегориальному варианту метода констант. Этот метод широко применяется в прикладных исследованиях, где обычно используются сложные стимулы, которые нетренированный испытуемый затрудняется классифицировать в терминах "больше" — "меньше", но хорошо понимает и может выполнить *задачу идентификации*, когда от него не требуется вынесения суждения только по одному из одновременно меняющихся сенсорных признаков при изменении физических параметров стимула. В качестве оценки чувствительности в этом методе используется полумежквартильный размах — Q(2). Однако эта оценка загрублена влиянием несенсорных факторов, приводящих к нестабильности критерия принятия испытуемым решения.

Для существенного уменьшения этого загрубления оценки Индлин (1979) предлагает ограничиваться в пределах одной непрерывной части опыта использованием одного сравниваемого стимула.

Методические рекомендации по выполнению учебных заданий по теме: "Локализация точки на шкале"

Задание 1. Определение величины иллюзии Мюллера—Лайера методом минимальных изменений

Цель задания: *Отработать метод минимальных изменений применительно к измерению разностного порога. Оценить величину иллюзии Мюллера—Лайера.*

Методика

Аппаратура. Задание отрабатывается на IBM-совместимом персональном компьютере. Для предъявления сигнала "Внимание" используются головные телефоны, соединенные со звуковым синтезатором персонального компьютера. Для выполнения учебного задания используется компьютерная программа muler.*exe*.

Стимуляция. На экране дисплея предъявляются на одной горизонтальной линии две стрелы: стандартный стимул (S_{st}) — стрела с наконечниками наружу, имеет длину 11 см и предъявляется всегда слева; и переменный стимул (S_{var}) — стрела с наконечниками внутрь и предъявляется всегда справа. Ее длина может меняться в пределах от 17 до 10 см. Время экспозиции стрел — 1 с.

Процедура опыта. При отработке задания каждый студент выступает сначала в роли испытуемого, а затем обрабатывает собственные экспериментальные данные. Испытуемый сидит на расстоянии 1 м до экрана дисплея. Каждая проба начинается с появления звукового сигнала “Внимание”, затем через 500 мс экспонируются стандартный и переменный стимулы (1 с). Следующая проба начинается через 2 с, в течение которых испытуемый должен дать свой ответ, нажимая на одну из 3-х клавиш на клавиатуре компьютера. Задача испытуемого заключается в том, чтобы сравнить переменный стимул со стандартным, используя три категории ответов: “меньше”, “равно” и “больше”. Для ответа могут использоваться следующие клавиши: <1>, <2>,<3> (на цифровой клавиатуре) или клавиши управления движением курсора — <←>, < ↓>,<→>.

Переменные стимулы предъявляются восходящими и нисходящими рядами, по 10 проб в каждом ряду. Всего 20 восходящих и 20 нисходящих рядов.

Обработка данных. После опыта студенту выдается компьютерная распечатка, в которой представлен полный протокол опыта, т.е. зафиксированы все ответы испытуемого на все стимулы (всего 400). Файл с полученными данными легко найти: его имя соответствует фамилии испытуемого, написанной латинскими буквами, а расширение — *mul*, например: *ivanov.mul.*

По данным протокола каждый студент должен вычислить следующие показатели:

1) нижний ($L_{l\uparrow}$ и $L_{l\downarrow}$) и верхний ($L_{h\uparrow}$ и $L_{h\downarrow}$) пороги в каждом ряду стимулов;

2) нижний (L_l) и верхний (L_h) пороги по опыту в целом (см. формулы (5) и (6)); оценить разброс полученных пороговых значений, рассчитав соответствующие значения стандартного отклонения — $\boldsymbol{\sigma}_l$ и $\boldsymbol{\sigma}_h$;

3) IU и DL (формулы (7) и (8));

4) PSE (формула (9));

5) количественно оценить по данным опыта выраженность иллюзии, рассчитав CE (формула (10).

Для выполнения необходимых статистических расчетов (среднее арифметическое и стандартное отклонение) следует воспользоваться статистическим пакетом “Stadia” (директория “Stadia”, командный файл — stadia.exe). После ввода полученных данных в электронную таблицу (40 значений L_l — в первую переменную и 40 значений L_h — во вторую) в меню статистических методов (F9) нужно выбрать самый первый пункт — “Описательная статистика” и указать номера анализируемых переменных: в нашем случае их две — 1, 2. После нажатия на клавишу <Enter> на экране распечатывается множество статистических показателей, в том числе — среднее арифметическое и стандартное отклонение для каждой переменной.

В качестве дополнительного задания по работе с данными можно рекомендовать построение графиков изменения пороговых значений в течение опыта, на котором в наглядной форме легко проанализировать возможные тенденции изменения верхнего и нижнего порогов, например: этап врабатывания, период стабилизации ответов и другие феномены динамики выполнения этой сенсорной задачи. Для этого, вернувшись в электронную таблицу (блок редактора данных), нужно нажать на клавишу F6 и перейти к построению нужного графика. В качестве типа графика выберите функциональный, далее укажите число графиков — 2 (один для L_l, другой для — L_h) и номера соответствующих им переменных (1 для первого графика и 2 — для второго); графики лучше нарисовать линиями, а разметку осей координат можно не делать.

Задание 2. Измерение порога частоты слияния мельканий методом установки

Цель задания. *Отработать метод установки применительно к измерению абсолютного порога. Оценить величину критической частоты слияния мельканий.*

Методика

Аппаратура. Задание выполняется на IBM-совместимом персональном компьютере, с параллельным портом которого соединен красный светодиод. Для выполнения учебного задания используется компьютерная программа *kcm.exe*.

Стимуляция. Светодиод расположен на расстоянии 1м от испытуемого. Мелькающий свет задается с помощью подачи на него коротких электрических импульсов длительностью 2 мс. Частота управляющих импульсов может изменяться испытуемым в диапазоне от 1 до 100 Гц с дискретностью 0.1 Гц.

Процедура опыта. При отработке задания каждый студент выступает сначала в роли испытуемого, а затем обрабатывает собственные экспериментальные данные. Испытуемый сидит на расстоянии 1 м до экрана дисплея. Задача испытуемого заключается в том, чтобы, регулируя частоту мелькания светодиода, установить пороговую (минимальную) величину частоты мелькания, при которой впервые появляется ощущение непрерывного свечения светодиода, т.е. отсутствие мельканий. Регулировка частоты мельканий осуществляется с помощью клавиш управления курсором на клавиатуре персонального компьютера: клавиша <←> служит для уменьшения частоты, а клавиша <→> — для ее увеличения. В инструкции отмечают, что испытуемый может свободно регулировать частоту, переходя через пороговое значение и снова возвращаясь назад.

Опыт состоит из 2-х серий — тренировочной и основной. В тренировочной серии испытуемому даются две попытки для установки порогового значения: начать из явно надпорогового диапазона стимулов (5 Гц), а затем — из явно подпорогового (60 Гц). При нахождении порога нужно нажать на клавишу “пробел”. В основной серии испытуемый осуществляет 20 установок: по 10 из надпорогового и подпорогового диапазонов. В каждой следующей пробе частота начального стимула меняется в случайном порядке.

Обработка данных. После опыта студенту выдается компьютерная распечатка, в которой представлен протокол опыта, где зафиксированы результаты всех 20 установок. При желании можно найти и файл данных: его имя соответствует фамилии испытуемого, а расширение — *kcm*.

По данным протокола каждый студент должен вычислить следующие показатели:

1) пороговую частоту слияния мельканий — RL, усреднив результаты всех 20 установок;

2) разброс пороговых значений, рассчитав стандартное отклонение — $\boldsymbol{\sigma}$;

3) статистическую ошибку, допущенную при вычислении RL (формула (4)).

Необходимые статистические вычисления производятся с помощью статистической системы “Stadia” в меню “Описательная статистика” (см. задание 1.).

Задание 3. Измерение порога различения длительности тональных сигналов методом постоянных раздражителей. Исследование влияния несенсорных факторов на пороговые меры

Цель *задания. 1. Практическая отработка метода на примере определения дифференциального порога при различных инструкциях испытуемому. 2. Освоение процедуры вычислений различных пороговых мер, получаемых в этом методе (интервал неопределенности, точка субъективного равенства, константная ошибка). 3. Сравнение зависимости различных пороговых мер от несенсорных факторов.*

Методика

Аппаратура. Задание отрабатывается на IBM-совместимом персональном компьютере. Для предъявления звуковых сигналов (тона частотой 1000 Гц) используются головные телефоны, соединенные со звуковым синтезатором персонального компьютера. Длительность стандартного стимула — 900 мс, длительность пяти переменных (сравниваемых) стимулов — 600, 750, 900, 1050 и 1200 мс.

Для выполнения учебного задания используется компьютерная программа *mc.exe*.

Процедура опыта. При отработке задания каждый студент выступает сначала в роли испытуемого, а затем обрабатывает собственные экспериментальные данные. Опыт состоит из одной тренировочной и трех основных серий. Испытуемому последовательно предъявляются два звуковых стимула, его задача — сравнить их по длительности. Место стандартного и сравниваемого стимулов в паре изменяется в квази-случайном порядке. Длительность сравниваемого стимула также меняется в квази-случайном порядке. Межстимульный интервал — 500 мс. Во время звучания каждого из стимулов на экране монитора последовательно появляются номера стимулов в паре (1 — 2, 1 — 2 и т.д.), что позволяет испытуемому определить, в какой момент времени нужно давать ответ. Если испытуемый не дал ответ в прошедшей пробе, то предъявление пары стимулов повторяется. Межпробный интервал, в течение которого испытуемому требуется дать ответ, используя трехкатегориальную систему ответов (“первый больше”, “равны” и “первый меньше”), равен 2 с. Для ответа используются цифровые клавиши — <1>, <2>, <3> .

Три основные серии отличаются друг от друга различными *инструкциями* испытуемому, что приводит к изменению его стратегии при выборе ответа. В значительной степени данный эксперимент повторяет известное исследование С. Фернбергера (цит. по: Бардин, 1976), в котором было четко показано, что при использовании в методе констант трехкатегорийной системы ответов значение дифференциального порога существенно зависит от сугубо *несенсорных* факторов.

Инструкция к *первой* серии[1] — нейтральная: испытуемый должен давать ответы согласно своим впечатлениям. Инструкция ко *второй* серии рассчитана на минимизацию ответов в промежуточной категории (“равно”): испытуемый должен пользоваться этой категорией в тех редких случаях, когда, несмотря на все усилия, он не способен различить сигналы по длительности. Инструкция к *третьей* серии рассчитана на максимизацию ответов “равно”: к этой категории необходимо относить все ощущения равенства или сомнения в сравнении стандарта и переменного стимулов; причем ответы “больше” и “меньше” нужно давать лишь при большой уверенности в разнице стимулов.

В каждой из этих серий испытуемым предъявляется 150 проб (по 30 предъявлений каждого переменного стимула). После каждой серии делается 2—3-минутный перерыв.

<u>Обработка результатов.</u> После окончания опыта испытуемый получает компьютерную распечатку, где для каждой серии приводятся частоты ответов “больше ”, “равно” и “меньше” для каждого из пяти переменных стимулов.

По каждой серии обработка результатов осуществляется следующим образом:

1. Строятся два графика с психометрическими кривыми в линейных координатах. На первом графике строятся психометрические кривые для 3-х категорий ответов (см. рис. 7—9). На втором — для 2-категориального варианта, при этом нейтральные ответы делятся поровну между пропорцией ответов “больше” и “меньше” (достаточно построить только одну кривую — для ответов “больше”).

---

[1] В том случае, если в эксперименте участвует группа студентов (9—15 человек), то имеет смысл учесть такой экспериментальный фактор, как эффект последовательности серий и, разбив группу на 3 подгруппы, для каждой из этих подгрупп определить свою последовательность прохождения серий эксперимента. После окончания эксперимента следует провести обработку результатов по группе в целом.

2. Для расчета всех пороговых показателей (IU, DL, PSE, CE) используется графический метод, основанный на способе линейной интерполяции.

3. Затем, используя 2-категориальный вариант расчетов, строятся психометрические функции в нормальных координатах для ответов “больше”.

С помощью метода наименьших квадратов по пяти экспериментальным точкам проводится наилучшая прямая, проходящая через эти точки. Для этого целесообразно воспользоваться статистическим пакетом “Stadia”. В редакторе данных длительности переменного стимула заносятся в первую переменную (это будут значения X), а z-оценки — во вторую (это будут значения Y). Затем переходят в меню статистических процедур (F9) и выбирают опцию “Простая регрессия (тренд)”. Войдя в нее, нужно указать номера переменных (1,2), а затем указать тип функции для построения регрессии — линейная. После этого программа построит для вас математическую модель ваших данных, представляя их в виде уравнения прямой: $Y = a_0 + a_1X$. Получив коэффициенты a_0 и a_1, можно без труда построить на графике аппроксимирующую прямую. Статистическая оценка адекватности сделанной линейной аппроксимации приводится внизу экрана результатов анализа. Не выходя из программы, можно легко вычислить и все необходимые показатели: PSE, $S_{\sigma+}$ и $S_{\sigma-}$. Это означает, что по уравнению регрессионной прямой нужно найти 3 неизвестных X по трем известным Y: z=0, z=+1 и z=-1. Для проведения расчетов нужно снова вернуться в меню статистических методов и, выбрав ту же опцию (“Линейная регрессия”), указать другой порядок переменных — 2,1. Это будет означать , что в качестве X мы выбираем z-оценки, а в качестве Y — длительность стимула. После расчета нового регрессионного уравнения нужно последовательно ввести три указанные выше величины z, и считать результат вычисления PSE, $S_{\sigma+}$ и $S_{\sigma-}$.

Далее вычисляются все необходимые пороговые показатели: IU, DL и CE.

4. Результаты обработки по каждой серии сводятся в итоговую таблицу:

Инстр-я	“Жесткая”			“Нейтральная”			“Либеральная”		
Обр-ка	3-х	2-х	норм.	3-х	2-х	норм.	3-х	2-х	норм.
IU									
DL									
PSE									
CE									
σ									

Таким образом в этой таблице в компактном виде должны быть представлены все полученные пороговые показатели в зависимости от инструкции (“жесткая”, “нейтральная” и “либеральная”) и метода обработки (3-х категориальный, 2-х категориальный и интерполяция в нормальных координатах).

Если проводится обработка групповых результатов (см. выше), то данные, полученные по трем группам испытуемых, усредняются и также сводятся в одну общую таблицу.

Обсуждение результатов. В ходе анализа полученных результатов следует оценить влияние такого мощного несенсорного фактора как инструкция на различные пороговые показатели, а также посмотреть, зависят ли рассчитанные показатели от метода обработки результатов.

В выводах нужно оценить возможности и ограничения метода констант применительно к задаче оценки сенсорной чувствительности.

Литература

Основная

1. *Бардин К.В.* Проблема порогов чувствительности и психофизические методы. М.: Наука, 1976. С. 69—278.

2. *Энген Т.* Психофизика 1. Различение и обнаружение // Проблемы и методы психофизики. / Под ред. А.Г.Асмолова, М.Б.Михалевской. М.: Изд-во Моск. ун-та, 1974.

Дополнительная

1. *Бардин К.В., Индлин Ю.А.* Начала субъективной психофизики: В 2 ч. М.: Изд. ИП РАН, 1993.

2. *Индин Ю.А.* Современные методы субъективной оценки различий в звучаниях // Обзорная информация НИКФИ. Вып.1(34). М., 1979. С.4—24.

3. *Михалевская М.Б., Скотникова И.Г.* Метод подравнивания: зависимость мер чувствительности от сенсорной задачи. Вес. Моск. ун-та. Сер. “Психология”. 1978. № 1. С.46-56.

4. *Guilford J.P.* Psychometric Methods. N.-Y.; Toronto; London: Mc-Grow-Hill, 1954. P. 86—153.

Приложение 1

Требования к оформлению отчета по учебному заданию

1. Отчет пишется на стандартных листах писчей бумаги. Все листы заполняются только с одной стороны. Номера листов проставляются в верхнем правом углу. Текст на листе ограничивается рамкой: сверху и снизу — 2—2.5 см, слева и справа — 2—2.5 см.

Каждый отчет начинается с *титульного листа*, который служит обложкой работы. Сверху на нем указывается принадлежность студента к учебному заведению, факультету, специализации или кафедре. В середине листа указывается название изучаемой темы или раздела и название учебного задания (например: “Общий психологический практикум: психологические измерения. Метод минимальных изменений”). Ниже и справа указывается фамилия и инициалы студента, номер академической группы, фамилия и инициалы преподавателя. Внизу титульного листа отмечают год выполнения работы.

Эта страница служит также для отметок преподавателя о выполнении учебного задания и замечаний по поводу подготовленного студентом отчета.

2. Структура отчета о выполнении учебного задания в практикуме:

Теоретическое введение и постановка проблемы (не более 3-х листов). В данном разделе отчета дается общая характеристика изучаемого метода, его характерных особенностей, даются определения необходимых терминов.

Формулировка цели и конкретных задач работы в соответствии с общей проблемой, рассмотренной в теоретическом введении (не более 0.5 листа).

Описание методики. В этот раздел входят следующие пункты: 1) сведения об испытуемом, дата и время проведения опыта; 2) описание использованной аппаратуры и программного обеспечения; 3) описание параметров стимуляции; 4) подробное описание процедуры опыта: какие стимулы предъявлялись, в каком порядке, какие ответы и в какой форме давал испытуемый; приводится инструкция испытуемому; если опыт состоял из нескольких серий, указывается их порядок. В том случае, если процедура опыта была нарушена, указывается причина.

Если опыт проводился на компьютере, следует указать имя файла результатов.

Результаты. В этой части необходимо описать полученные данные, методы их обработки и привести основные результаты. Если использовались нестандартные способы обработки результатов, то их описанию стоит уделить особое внимание. Если использовались методы статистического анализа, то необходимо привести соответствующие формулы или сослаться на использованный статистический пакет при работе на компьютере. В последнем случае, как правило, следует привести стандартную распечатку полученных результатов обработки.

Итоговые результаты проведенных измерений сводятся в одну таблицу. Каждая таблица должна быть пронумерована и иметь соответствующее название, где нужно четко выразить основное содержание данной таблицы.

Рисунки, иллюстрирующие основное содержание работы, должны быть также пронумерованы и начерчены на координатной бумаге. В том случае, если по рисунку производится вычисление каких-либо результатов (как, например, в случае с психометрической функцией), то следует обратить внимание на выбор подходящего масштаба. Если рисунок делается с помощью компьютерной программы, то стоит позаботиться о введении координатной сетки. На графиках должны быть указаны все параметры, необходимые для однозначного понимания графика. Подрисуночные подписи и обозначения на графиках должны давать полную информацию, чтобы не возникала необходимость для понимания графика обращаться к тексту отчета.

Рисунки и таблицы рекомендуется выполнять на отдельных листах.

Обсуждение результатов и выводы должны соответствовать целям и задачам работы. В том случае, если получен нестандартный и неожидаемый результат, то безусловно следует уделить особое внимание его интерпретации и попытаться объяснить причины его появления. Если работа выполнялась в рамках какой-либо модели, то следует сделать четкое заключение о соответствии полученных результатов ее предположениям.

Выводы должны быть короткими и конкретными.

Литературные ссылки оформляются в соответствии с требованиями, предъявляемыми ГОСТом к научным статьям.

Таблица для перевода значений p в значения z

p	0,01	0,02	0.03	0,04	0,05	0,06	0,07	0,08	0,09	0,10
Z	*-2,33*	*-2,05*	*-1,88*	*-1,75*	*-1,64*	*-1,55*	*-1,48*	*-1,41*	*-1,34*	*-1,28*
p	0,11	0,12	0,13	0,14	0,15	0,16	0,17	0,18	0,19	0,20
Z	*-1,23*	*-1,18*	*-1,13*	*-1,08*	*-1,04*	*-0,99*	*-0,95*	*-0,92*	*-0,88*	*-0,84*
p	0,21	0,22	0,23	0,24	0,25	0,26	0,27	0,28	0,29	0,30
Z	*-0,81*	*0,77*	*0,74*	*0,71*	*0,67*	*0,64*	*0,61*	*0,58*	*0,55*	*0,52*
p	0,31	0,32	0,33	0,34	0,35	0,36	0,37	0,38	0,39	0,40
Z	*-0,50*	*-0,47*	*-0,44*	*-0,41*	*-0,39*	*-0,36*	*-0,33*	*-0,31*	*-0,28*	*-0,25*
p	0,41	0,42	0,43	0,44	0,45	0,46	0,47	0,48	0,49	0,50
Z	*-0,23*	*-0,20*	*-0,18*	*-0,15*	*-0,13*	*-0,10*	*-0,08*	*-0,05*	*-0,03*	*-0,00*
p	0,51	0,52	0,53	0,54	0,55	0,56	0,57	0,58	0,59	0,60
Z	*+0,03*	*+0,05*	*+0,08*	*+0,10*	*+0,13*	*+0,15*	*+0,18*	*+0,20*	*+0,23*	*+0,25*
p	0,61	0,62	0,63	0,64	0,65	0,66	0,67	0,68	0,69	0,70
Z	*+0,28*	*+0,31*	*+0,33*	*+0,36*	*+0,39*	*+0,41*	*+0,44*	*+0,47*	*+0,50*	*+0,52*
p	0,71	0,72	0,73	0,74	0,75	0,76	0,77	0,78	0,79	0,80
Z	*+0,55*	*+0,58*	*+0,61*	*+0,64*	*+0,67*	*+0,71*	*+0,74*	*+0,77*	*+0,81*	*+0,84*
p	0,81	0,82	0,83	0,84	0,85	0,86	0,87	0,88	0,89	0,90
Z	*+0,88*	*+0,92*	*+0,95*	*+0,99*	*+1,04*	*+1,08*	*+1,13*	*+1,18*	*+1,23*	*+1,28*
p	0,91	0,92	0,93	0,94	0,95	0,96	0,97	0,98	0,99	0,995
Z	*+1,34*	*+1,41*	*+1,48*	*+1,55*	*+1,64*	*+1,75*	*+1,88*	*+2,05*	*+2,33*	*+2,58*

Глава 2. МЕТОДЫ ОБНАРУЖЕНИЯ СИГНАЛА

§ 1. Общие понятия

В этой главе рассматриваются методы, отличающиеся от предыдущей группы методов новым подходом к локализации точки на психологической шкале, иначе говоря, другим подходом к измерению граничного шкального значения, разделяющего имеющееся множество стимулов на два класса: обнаруживаемые и необнаруживаемые, различаемые и неразличаемые и т.п.

В классических психофизических методах, хотя и изучаются сенсорные способности наблюдателя, не ставится вопрос о *вероятности обнаружения стимула,* а учитывается лишь вероятность ответов испытуемого “Да” (слышу или вижу). Однако легко себе представить такую ситуацию, когда испытуемый, находясь в ситуации тестирования (экспертизы), захочет показать максимум своих сенсорных способностей, и будет давать ответ “Да” почти в каждой пробе. Естественно, что в таком случае количество утвердительных ответов не будет сколько-нибудь точно отражать его предельные сенсорные способности. Надежда психолога-эксперта на честность испытуемого, по-видимому, не самое лучшее средство для обеспечения надежности проводимых измерений. Таким образом, достаточно очевидно, что результат пороговых измерений может сильно зависеть от *стратегии* испытуемого давать ответы определенного рода, и, следовательно, появляется задача прямого учета поведения наблюдателя в ситуации принятия решения об обнаружении или различении сигнала.

Новая методология, называемая *психофизической теорией обнаружения сигнала* (Green , Swets , 1966), содержит в себе представление о наблюдателе как не о пассивном приемнике стимульной информации, но как об активном субъекте принятия решения в ситуации неопределенности.

Вкратце этот подход можно охарактеризовать следующим образом. В стимульном потоке выделяется та его часть, на которую указанием ее пространственной и/или временной области или ее характерного паттерна обращается внимание *наблюдателя*. Эта выделенная часть называется *стимулом* или предъявлением (стимула). Выделяется некоторый физический признак (свойство, характеристика стимульного потока), который может присутствовать в одних пробах — *значащий или сигнальный стимул,* и отсутствовать в других — *пустой стимул*. Наблюдатель, от которого требуется *обнаруживать этот признак*, решает задачу бинарной классификации: относит каждое предъявление к одному из двух классов — "Нет признака", "Есть признак". Эта задача решается путем установления *схемы соответствия* (которая называется также *правилом принятия решения*) между особенностями сенсорного образа предъявляемого стимула и выбираемым решением. Эта схема соответствия может корректироваться под влиянием как предварительного информирования наблюдателя о частоте сигнальных или пустых стимулов в последующих предъявлениях, так и обратной связи — оценки правильности принимаемых наблюдателем решений.

В следующих трех разделах будут описаны три классических метода обнаружения сигналов: метод "Да-Нет", двухальтернативный вынужденный выбор и метод оценки уверенности.

§2. Метод "Да-Нет"

В этом методе используются два стимула: один значащий — $<S>$, и другой пустой — $<N>$. Предъявления следуют друг за другом обыкновенно через более или менее регулярные интервалы времени и после каждого предъявления испытуемый отвечает "Да", если был сигнал, или "Нет", если он не обнаружил сигнала. Предъявление стимулов полностью *рандомизировано*, т.е. каждое очередное предъявление независимо от предыдущих может может быть с некоторой вероятностью P(S) сигнальным (и, следовательно, с вероятностью P(N) = 1 - P(S) — пустым); P(S) и P(N) сохранятся постоянными на протяжении всей серии предъявлений. Таким образом, если общее число предъявлений N в эксперименте достаточно велико, то число сигнальных и пустых предъявлений приблизительно равно, соответственно $N \cdot P(S)$ и $N \cdot P(N)$ (очевидно, $N \cdot P(S) + N \cdot P(N) = N$).

Рассмотрим теперь возможные комбинации <предъявление — ответ>, которые могут встретиться в эксперименте. Их четыре: <S — “Да”>, <N — “Нет”>, <S — “Нет”>, <N — “Да”>, причем первые два сочетания являются правильными, два последние — ошибочными исходами. Каждое их этих сочетаний имеет свое специальное название, как это показано в табл. 1.

Таблица 1

Исходы эксперимента по обнаружению сигнала

Стимул	Ответ *Да*	Ответ *Нет*
<*S*>	H	O
<*N*>	FA	CR

Попадание и ложная тревога будут в дальнейшем обозначаться через H (от английского hit) и FA (от английского false alarm). Обозначения для пропусков и правильных отрицаний — O (omission) и CR (correct rejection). Пусть мы пересчитали количество сочетаний каждого типа: n (H), n (FA), n (O), n (CR). Очевидно, что:

$$n(H) + n(O) = N \cdot P(S), \qquad (1)$$
$$n(FA) + n(CR) = N \cdot P(N) \quad , \qquad (2)$$

Зная эти качества и нормировав каждое из них по N (т.е., поделив на общее количество предъявленных проб), мы получим статистические оценки вероятностей появления исходов каждого типа:

$$P(H) = n(H)/N,\ P(O) = n(O)/N, \ldots \text{ и т.д.} \qquad (3)$$

Однако такие вероятности еще не говорят нам прямо о способности наблюдателя обнаруживать сигнал. Действительно, величина p(H) зависит не только от того, как часто наблюдатель идентифицирует <S> как сигнал, но и от того, сколь часто предъявлялось в эксперименте <S>. Поэтому, чтобы охарактеризовать деятельность испытуемого в данном эксперименте, отделив ее от деятельности экспериментатора (решающего, в частности, сколько раз предъявить <S>, а сколько — <N>), принято представлять результаты эксперимента в виде оценок *условных вероятностей* — вероятностей того, что испытуемый ответит правильно (неправильно) при условии, что был предъявлен данный стимул. Такие вероятности обозначаются так: P ("Да"/S), P ("Да"/N), P ("Нет"/S), P ("Нет"/N). В частности, первая из этих вероятностей есть вероятность правильного ответа при условии, что было предъявлено <S>. Легко видеть, что:

$$P(\text{"Да"}/S) = P(H)/P(S) = n(H)/\, N \cdot P(S), \qquad (4)$$
$$P(\text{"Да"}/N) = P(FA)/P(N) = n(FA)/\, N \cdot P(N). \qquad (5)$$

Если вычислены две эти условные вероятности, вычисление двух остальных уже не требуется. Они не несут дополнительной информации, т.к.:

$$P(\text{"Нет"}/S) + P(\text{"Да"}/S) = 1, \qquad (6)$$
$$P(\text{"Нет"}/N) + P(\text{"Да"}/N) = 1. \qquad (7)$$

Итак, при данных (выбранных экспериментатором) величинах N и P(S) результаты эксперимента обычно представляют только двумя условными вероятностями: вероятностью попадания — p(H)=P("Да"/S) и вероятностью ложной тревоги p(FA)=P("Да"/N).

Заметим, что при всех приведенных выше расчетах из общего числа N предъявлений обычно исключают несколько первых (порядка 40—50), предполагая, что в этих первых пробах испытуемый постоянно меняет схему соответствия, "подстраивая" ее к информации, полученной от экспериментатора и в ходе эксперимента. Когда схема соответствия устанавливается стабильно, говорят, что решение задачи вышло на *асимптотический уровень*. Асимптотический уровень характеризуется тем, что если все число предъявлений (после исключения первых) произвольно разбито на несколько групп и по каждой из них в отдельности вычислено P(H) и P(FA), то все эти пары не будут статистически значимо отличаться друг от друга.

Полная характеристика эксперимента требует указания еще двух факторов: наличие/отсутствие предварительной информации и наличие/отсутствие обратной связи. *Предварительная информация* — это формальный признак, означающий сообщение испытуемому величины P(S). Например: “В 80% всех проб будет предъявляться пустой стимул” (т.е. P(S) = 0,2) или — “Сигнальное предъявление будет встречаться в 3 раза чаще пустого” (P(S)/P(N) = 3, т.е. P(S) = 0,75). Сама инструкция, разъяснение испытуемому формы предъявления, характера сигнала и т.п. — все это не входит в термин “предварительная информация”. Заметим, что предварительная информация, если она вводится, может быть и ложной, т.е. испытуемому может сообщаться не та величина P(S), которая есть на самом деле. Эта специальная модификация “Да-Нет”-эксперимента, которая здесь рассматриваться не будет. Термин *обратная связь* включает информацию об истинности/ложности ответов испытуемого, сообщаемую ему после каждого предъявления, или сообщение о частоте правильных ответов, даваемое после некоторой группы (скажем, через каждые 50) предъявлений. В специальных модификациях метода такая обратная связь также не всегда должна быть истинной. Иногда, например, используют такой вариант, когда после каждой пробы (предъявления) испытуемому с вероятностью P(k) сообщают важную информацию о правильности или ложности его ответа, а с вероятностью 1 - P(k) его “обманывают” (в этом варианте P(k) — формальная мера правдивости обратной связи).

Цель введения обратной связи и предварительной информации — попытка контроля схемы соответствия между свойствами ощущений и принимаемыми решениями, которую устанавливает испытуемый (правила принятия решения). Очевидно, однако, что если испытуемый не очень заинтересован в том, чтобы почаще отвечать правильно, то такой контроль может оказаться неэффективным[1]. Кроме того, испытуемый может, устанавливая правило принятия решения, руководствоваться неизвестными экспериментатору субъективными "весами" различных типов ошибок. Например, он может стараться минимизировать число пропусков и не очень заботиться об уменьшении числа ложных тревог (т.е. "цена" пропуска выше "цены" ложной тревоги). Чтобы сделать контроль правила принятия решения более эффективным и дифференцированным, обратная связь может быть дополнена *системой "выплат" и "штрафов",* соответственно за верные и ложные ответы, организованной в денежной или какой-либо другой (например, просто игровой) форме. Это можно записать в форме следующей *платежной матрицы*:

	Да	Нет
Сигнал	V	-V
Шум	-W	W

где V и W — положительные числа. Такая форма представления особенно удобна, так как она позволяет ограничиться только двумя числами, V и W, для характеристики всей платежной матрицы. Матрица называется симметричной, если V = W. Для определения оптимального правила принятия решения, т.е. такого из имеющегося у наблюдателя набора возможностей, которое максимизирует выигрыши, решающее значение имеет соотношение не самих V и W, а $P(S) \cdot V$ и $P(N) \cdot W$ (они совпадают, только если $P(S) = 0,5$). Если $P(S) \cdot V = P(N) \cdot W$, правило принятия решения должно быть установлено так, чтобы

[1] Дополнительная информация о критериях оптимальности принимаемого наблюдателем решения приведена в приложении к данной главе.

минимизировать вероятности ошибок обоих родов. Если же $P(S)\cdot V > P(N)\cdot W$, то правило целесообразно изменить так, чтобы сделать возможно меньшей вероятность пропуска, $1\text{-}p(H)$, даже если при этом увеличивается вероятность ложной тревоги, $p(FA)$.

Возникает вопрос: что ограничивает набор возможных схем соответствия? Почему, в частности, испытуемый не всегда может выработать “правильную” схему соответствия, при которой $p(H)=1$ и $p(FA)=0$? Ответ на эти вопросы требует построения формальной модели следующих процессов: 1) какое соответствие существует между предъявлениями <S> и <N> и их сенсорными репрезентациями; 2) как по данной сенсорной репрезентации строится ответ (“Да” или “Нет”). Мы изложим здесь одну из простейших моделей, отвечающих на эти вопросы.

Суть модели состоит в следующем. Любой стимул (<S> или <N>) связан с его сенсорными репрезентациями *стохастически* (т.е. вероятностно, случайно), а не детерминистически. Это означает, что один и тот же стимул, повторяясь в различных пробах, вызывает различные сенсорные образы, так что в каждой отдельной пробе можно говорить только о вероятностях возникновения тех или иных сенсорных образов. Причины такой стохастичности многочисленны. С одной стороны, они могут лежать в природе самого стимула (например, количество квантов, излучаемых источником света в данном направлении в единицу времени — величина принципиально стохастическая) и в ограниченной точности приборных измерений. С другой стороны, стохастичность связана со случайными флуктуациями в сенсорной системе, например, со спонтанной нервной активностью в проводящих путях. Последняя, в частности, обеспечивает наличие различных сенсорных образов и в том случае, если пустой стимул <N> представляет собой отсутствие энергии в данной пространственно-временной области. Кроме того, известный вклад в стохастичность сенсорных эффектов безусловно вносят и так называемые внешние факторы: нестабильность стимуляционной аппаратуры, различного рода помехи и т.д.

Далее в излагаемой модели предполагается, что установленное правило соответствия имеет *детерминистическую* структуру, т.е. данный сенсорный образ, если он в точности повторился в двух пробах (причем за время между пробами схема соответствия не изменилась), вызовет всегда один и тот же ответ. Другими словами, любое правило принятия решения *однозначно* разбивает множество всех возможных ощущений на два класса — один, связанный с ответом “Да”, другой — с ответом “Нет”. На рис. 1 точками заполнены те области, которые связаны с ответом “Да”. На рис. 2 области с горизонтальной (вертикальной) штриховкой соответствуют ощущениям, которые могут быть вызваны стимулами <S> и <N>.

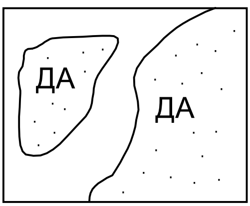


Рис.1. Два множества ощущений, вызывающих ответ "Да"

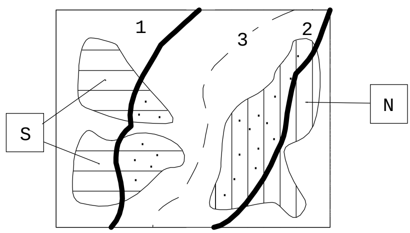


Рис.2. Множества непересекающихся ощущений, вызванных значащим и пустым стимулами: S - значащий стимул; N - пустой стимул; 1,2 и 3 - линии, показывающие границы разбиения множества ощущений

Линии 1, 2 и 3 показывают границы разбиения, соответствующего трем схемам соответствия, причем область “Да” при всех схемах соответствия лежит слева от этих границ. Рассмотрим сначала схему соответствия 1. Мы видим, что при такой схеме <N> всегда будет идентифицироваться правильно, т.е. p(FA)=0. Однако, <S> иногда (когда ощущение, вызванное <S>, попадает правее границы — эта область помечена точками) вызовет ответ “Нет”, т.е. испытуемый будет иногда пропускать сигнал, p(H)<1. При схеме соответствия 2 ситуация обратна. Испытуемый всегда идентифицирует <S> как “Да”, т.е. p(H)=1, но иногда (эта область помечена точками) <N> будет вызывать ответ “Да” (ложная тревога), т.е. p(FA)>0. Легко видеть, однако, что при данном сорасположении ощущений, вызываемых <S> и <N>, испытуемый может в принципе выработать такую схему соответствия (граница 3, штриховая линия), при которой можно избежать ошибок, т.е. p(FA)=0 и p(H)=1. Причина, обеспечивающая эту возможность, заключается в том, что указанные области не пересекаются, т.е. нет ни одного ощущения, которое могло бы быть вызвано (пусть с разной вероятностью) как <S>, так и <N>. Если это условие не выполняется (см. рис. 3), то, очевидно, при любой схеме соответствия испытуемый будет совершать ошибки того или иного рода (O или FA), либо и те, и другие.

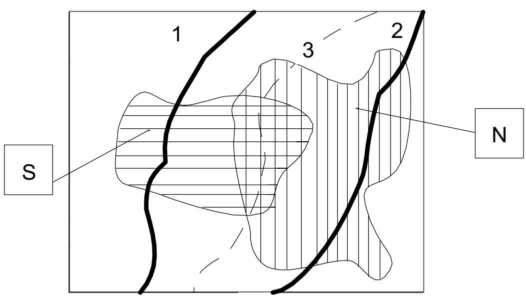


Рис. 3. Два пересекающихся множества ощущений, вызванных значащим и пустым стимулами

Такова суть модели. Чтобы представить модель в количественной форме, допускаются *два дополнительных упрощения.* Первое из них может быть разъяснено следующим образом. Схема соответствия с содержательной точки зрения представляет собой соответствие данного ответа некоторому комплексу свойств сенсорного образа: “Если образ обладает свойствами таким-то и таким-то, то следует выбрать ответ “Да”, в противном случае — “Нет”. Очевидно, что не все свойства образа при этом используются. Рассматриваемое упрощение состоит в предположении, что решение принимается всегда на основе интенсивности какого-то *одного* качества сенсорных образов (“сладкость”, “наклонность”, “яркость” и т.п.), причем правило принятия решения имеет форму: “Если интенсивность (выраженность) качества больше некоторой величины *С*, то следует выбрать “Да”, в противном случае — “Нет”. Интенсивность качества, как это видно из предыдущей фразы, предполагается представимой *действительным числом*. Таким образом все возможные значения интенсивности данного качества занимают какую-то часть оси действительных чисел (например, всю положительную полуось), причем каждое из этих значений при предъявлении данного стимула может быть вызвано с тем или иным *правдоподобием*. Если значения интенсивности сенсорных образов образуют *непрерывный континуум*, то это правдоподобие выражается не вероятностью, а *плотностью вероятности*. Плотность вероятности возникновения ощущения со значением интенсивности ощущения X при подаче стимула A условимся обозначать через f (X/A).

Вернемся теперь к нашей ситуации, где стимул есть либо <S>, либо <N>. Каждому из стимулов соответствует своя функция плотности вероятности: f (X/S) и f (X/N) (рис. 4).

Согласно принятому утверждению, правило принятия решения определяется выбором *граничной точки С* (ее еще называют критической точкой или величиной *критерия принятия решения* о наличии сигнала), такой, что если интенсивность X в данной пробе превышает С, то следует ответ “Да”, если же не превышает, то — “Нет”. Легко видеть по рисунку, что вероятность ложной тревоги p(FA) равна вероятности того, что интенсивность X (при условии, что предъявлен <N>) превзойдет С, т.е. равна заштрихованной области под кривой f (X/N). Вероятность попадания p(H) равна

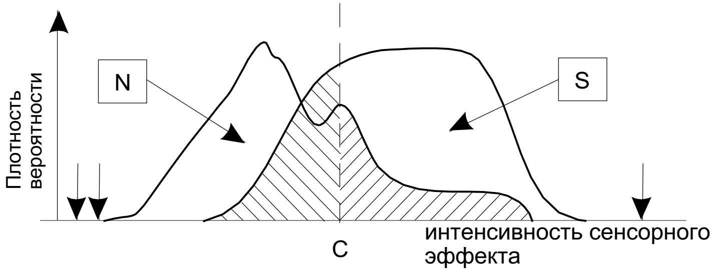


Рис. 4. Общая модель обнаружения сигнала: справа – распределение сенсорных эффектов при воздействии значащего стимула, слева – пустого стимула

вероятности того, что X (при условии, что предъявлен <S>) превзойдет C, т.е. равна незаштрихованной области под кривой f (X/S).

$$p(H) = \int_{c}^{\infty} f(x / S) / dx \qquad (8)$$

$$p(FA) = \int_{c}^{\infty} f(x / N) / dx \qquad (9)$$

Если критерий C находится далеко вправо (показано на рис. 4 одной стрелкой), то, очевидно, p(FA)=p(H)=0. Если теперь начать двигать критерий справа налево, то при каждом очередном значении мы будем получать новую пару p(FA) и p(H), причем оба значения будут возрастать (или по крайней мере не убывать), пока при достаточно далеком левом положении C оба не станут равны 1 (показано двумя стрелками на рис. 4). Поскольку каждое значение C однозначно определяет пару чисел p(FA) и p(H), то ему можно поставить в соответствие точку внутри квадрата (рис. 5), на вертикальной стороне которого откладывается p(H), а на горизонтальной — p(FA), и таким образом представить результат работы наблюдателя .

Полученная по этим точкам кривая называется *рабочей характеристикой наблюдателя* или просто — *РХ*. Лю-

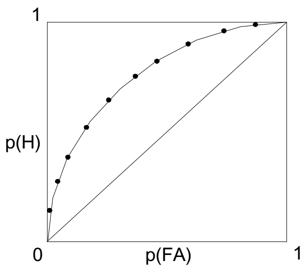


Рис.5. Рабочая характеристика идеального наблюдателя

бая пара распределений, f(X/S) и f(X/N) однозначно определяет РХ, но обратное неверно: одна и та же РХ может определяться различными парами f(X/S) и f(X/N). РХ идет из точки (0,0) квадрата в точку (1,1) и при этом располагается выше его главной диагонали. Последнее следует из того, что распределение f(X/S) сдвинуто вправо относительно f(X/N), т.е. p(H) превышает p(FA)[1].

Вероятности p(H) и p(FA) меняются *содружественно*, т.е. нельзя только путем изменения схемы соответствия одновременно увеличить

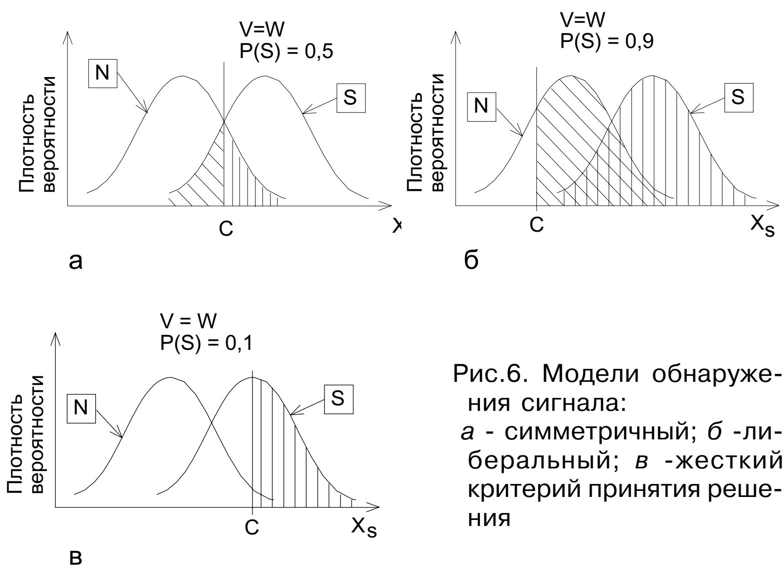


Рис.6. Модели обнаружения сигнала:
а - симметричный; *б* -либеральный; *в* -жесткий критерий принятия решения

[1] Очевидно, что когда наблюдатель фактически не обнаруживает сигнал и p(H)=p(FA), точки РХ будут лежать на диагонали.

одну из них и уменьшить другую (или, что то же самое, нельзя одновременно уменьшить или увеличить вероятности ошибок обоих родов, FA и O). Это очень важное положение верно для любых пар f(X/S) и f(X/N). Из него следует, что только пара этих вероятностей, а не каждая в отдельности, характеризует сенсорную способность наблюдателя.

Допустим, в эксперименте с симметричной платежной матрицей (V=W) и P(S) = 0,5) испытуемый установил положение критерия, как это показано на рис. 6а.

Результаты этого мысленного эксперимента с так называемым *симметричным критерием* представлены в таблице 3.

Таблица 2

Вероятности исходов эксперимента с симметричной платежной матрицей и P(S)=0.5

Стимул	Ответ	
	Да	*Нет*
<*S*>	0.75	0.25
<*N*>	0.25	0.75

Это положение критерия *оптимально* в том смысле, что суммарный выигрыш испытуемого в этом случае будет максимален.

Пусть теперь в следующем эксперименте платежная матрица осталась симметричной, а P(S)=0.9.

Таблица 3

Вероятности исходов эксперимента с симметричной платежной матрицей и P(S)=0.9

Стимул	Ответ	
	Да	*Нет*
<*S*>	0.95	0.05
<*N*>	0.60	0.40

Теперь (рис. 6б), чтобы сохранить тот же выигрыш, наблюдателю необходимо сдвинуть критерий так, чтобы p(H) резко возросло, даже за счет возрастания p(FA) — теперь важнее не пропустить сигнал, чем не дать ложную тревогу! Следовательно, критерий *C* сдвинется влево. В данном случае говорят, что наблюдатель использует *либеральный* критерий.

Пусть в третьем эксперименте при симметричной платежной матрице P(S) установили равной 0.1.

Таблица 4

Вероятности исходов эксперимента с симметричной платежной матрицей и P(S)=0.1

Стимул	Ответ	
	Да	*Нет*
<*S*>	0.65	0.35
<*N*>	0.94	0.06

В этой ситуации (рис. 6в) критерий должен быть сдвинут вправо, и в этом случае говорят об использовании *строгого* критерия. Аналогичные изменения положения критерия принятия решения можно рассмотреть и при изменениях платежной матрицы при постоянной величине P(S).

Для каждой пары f(X/S) и f(X/N), если заданы V,W и P(S), может быть рассчитано оптимальное положение *C* — то, при котором выигрыш максимален. В соответствии с данной логикой можно исследовать вопрос, насколько реальное положение критерия, выбираемое испытуемым, близко к оптимальному. Но, разумеется, это можно сделать лишь в том случае, если мы можем восстановить по результатам экспериментов теоретическую схему, т.е. построить функции распределения f(X/S) и f(X/N) и найти критерий *C*.

Итак, перед нами стоит задача *восстановления теоретической схемы по экспериментальным данным*. Прежде всего, разберемся в том, что представляют собой экспериментальные данные. Пусть выбраны стимулы <S> и

<N> и проведен эксперимент по методу “Да-Нет”. Результатом эксперимента является пара вероятностей p(H), p(FA). Далее какие-то параметры эксперимента меняются (изменятеся P(S) и/или платежная матрица, или снимается обратная связь и заменяется на предварительную информацию или что-то еще), и эксперимент повторяется с теми же <S> и <N>. Получаем, вообще говоря, другую пару p(H), p(FA). Повторяя эксперимент несколько раз, мы будем иметь в результате несколько пар p(H), p(FA), т.е. несколько точек РХ. Разумеется, и это очень важно, мы можем считать все эти пары p(H) и p(FA) точками одной РХ лишь постольку, поскольку предполагается, что изменения экспериментальных параметров могут привести *только* к изменению положения критерия *С*, но не к изменению схемы соответствия, в более широком смысле слова включающем возможное привлечение новых сенсорных качеств, замену одного качества на другое и в результате, если это новое качество одномерно, — получению новой пары распределений f(X/S) и f(X/N). Таким образом, проблема формулируется так: по нескольким точкам РХ нужно восстановить f(X/S), f(X/N) и С. Однако, мы уже говорили, что в таком виде проблема не решается, так как даже если бы была известна вся РХ (т.е. все точки, а не несколько, чего никогда, естественно, не бывает), распределения f(X/S) и f(X/N) не восстановимы однозначно. Поэтому в модели, которую мы излагаем (обычно называемой, хотя и не совсем точно, *теорией обнаружения сигналов*, ТОС) принимается еще одно упрощающее предположение (впрочем, в отличие от первого, оно допускает прямую экспериментальную проверку, о чем речь пойдет ниже): существует такая монотонная трансформация оси интенсивности, в результате которой оба распределения становятся *нормальными*. Для краткости трансформированную ось мы будем обозначать просто через z и говорить о z-значениях. Под монотонной трансформацией понимается система всевозможных растяжений и сжатий различных областей оси X так, что если точка q лежит левее r, то после трансформации это отношение сохраняется.

Примером такой трансформации является логарифмирование, растягивающее положительную полуось действительных чисел на всю действительную ось. Итак, мы имеем два нормальных распределения, причем всегда можно считать, что на оси выбрана такая позиция нуля и такой масштаб, что f(Z/N) имеет центр в нуле и стандартное отклонение, равное 1. Для восстановления теоретической картины, таким образом, необходимо определить положение центра и стандартное отклонение распределения f(Z/S).

Если допустить, что $\sigma_{s,n} = 1$, т.е. дисперсии обоих распределений равны, а центр распределения f(Z/S) сдвинут вправо от центра распределения f(Z/N) на величину *a*, тогда

$$f(z/S) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-a)^2}{2}}. \qquad (10)$$

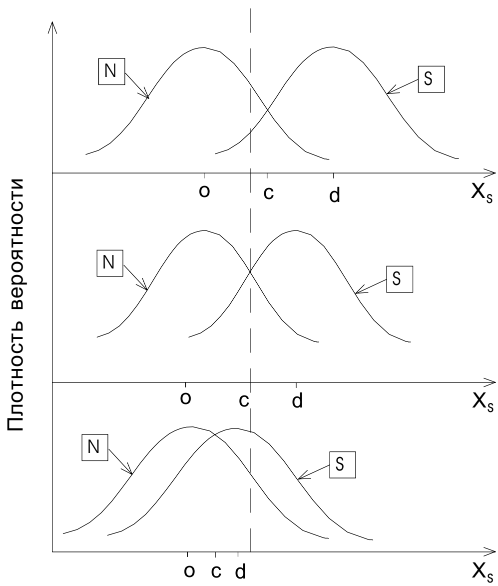


Рис. 7. Модель ТОС при различных уровнях обнаружимости сигнала

В этом случае вместо *a* обыкновенно пишут специальный символ d' и называют эту величину мерой чувствительности наблюдателя к сигналу. Чувствительность к сигналу характеризуется степенью отличия Z-величин, вызываемых <S>, от Z-величин, вызываемых <N>. Чем меньше величина d', тем больше перекрываются области Z-значений, соответствующих <S> и <N> (рис. 7).

Легко видеть, что при одном и том же положении критерия *C*, а следовательно, при одной и той же величине p(FA), величина p(H) тем ближе к p(FA), чем меньше d'. Если d' = 0, то p(FA) = p(H) при всех *C* и, следовательно, РХ в таком эксперименте совпадает с главной диагональю квадрата (рис. 8). Если d' > 0, РХ лежит выше диагонали и имеет гладкий и симметричный вид относительно побочной диагонали, идущей из (0,1) в (1,0). Чем больше d', тем более выпукла РХ влево-вверх и тем дальше она отстоит от главной диагонали. Как же практически вычислить d' и *C* по результатам эксперимента? Сколько точек РХ следует для этого иметь?

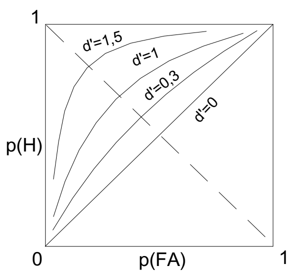


Рис.8. РХ при различных уровнях обнаружимости стимула

Оказывается, достаточно только одной точки, т.е. только одной пары p(FA), p(H). Действительно,

$$p(FA) = \frac{1}{\sqrt{2\pi}} \int_{c}^{\infty} e^{-z^2/2} dx \,. \qquad (11)$$

Это уравнение необходимо решить относительно *C*. Введем новый термин: нахождение *C* по P в уравнении (12):

$$P = \frac{1}{\sqrt{2\pi}} \int_{c}^{\infty} e^{-z^2/2} dx \qquad (12)$$

называется Z-преобразованием P:

$$C = Z\,[P]. \qquad (13)$$

Сделать Z-преобразование можно по обычной таблице нормального распределения. Если есть таблица, показывающая для каждого *C* значение интеграла (12), то нужно попросту отыскать в таблице значение интеграла, наиболее близкое к P, и посмотреть слева, какому C оно соответствует. Легко показать, что уравнение (11) в терминах Z-преобразования имеет решение:

$$C = -Z\,[p(FA)]. \qquad (14)$$

Теперь допустим, что *C* найдено. Как, зная p(H), найти величину d'? Рассмотрим теоретическую картинку, из которой удалено распределение, соответствующее N (оно уже не понадобится, см. рис. 9а). Сдвинем все распределение вдоль оси Z вместе с критерием *C* влево так, чтобы центр совместился с точкой 0. Критерий *C* при этом, очевидно, займет позицию (C - d'), а заштрихованная область не изменится и останется равной по площади p(H) (см. рис. 9б). Но наше сдвинутое распределение имеет центр в нуле и единичную дисперсию.

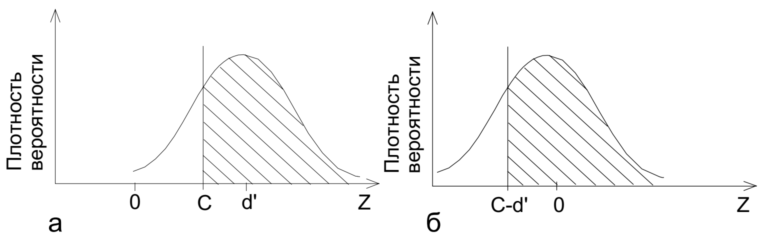


Рис.9. Теоретическое распределение ощущений при действии значащего стимула:
а - .сдвинутое на величину d' относительно "шумового" распределения; б - с центром, смещенным влево до точки 0;
ось X - величина единичного среднеквадратичного отклонения; ось Y - плотность вероятности величины сенсорного эффекта; точка C - положение критерия.

Следовательно:

$$p(H) = \frac{1}{\sqrt{2\pi}} \int\limits_{c-\mathrm{d}}^{\infty} e^{-z^2/2} dx \quad \text{и} \tag{15}$$

$$d' - C = z[p(H)]. \tag{16}$$

Сопоставив (14) и (16), получим:

$$d' = z[p(H)] - z[p(FA)]. \tag{17}$$

Допустим теперь, что проведен новый эксперимент с измененными параметрами, так что получена новая пара p(FA) и p(H). Если наше предположение относительно f(Z/S) и f(Z/N) верно (т.е. они оба нормальны и имеют одну и ту же дисперсию), то, несмотря на изменение величины *C* прямо определяемой по формуле (14), величина d', определяемая по формуле (17), должна оставаться постоянной. Мы приходим к важному заключению: если по оси абсцисс откладывать величины Z[p(FA)], а по оси ординат — z[p(H)], то точки РХ должны выстроиться в прямую линию, описываемую уравнением (17): z[p(H)] = =z[p(FA)] + d', и наклоненную под 45° к оси абсцисс. График зависимости Z[p(H)] от Z[p(FA)] (см. рис. 10) называется РХ в двойных нормальных координатах. Из соотношения (17) вытекает способ экспериментальной проверки предположений, принятых о нормальности распределений и равенстве дисперсий. Пусть мы провели K экспериментов и получили K точек РХ (K ≥2).

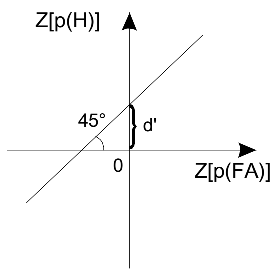


Рис.10. РХ в двойных нормальных координатах, $\sigma_S=\sigma_N$.

Построим РХ в двойных нормальных координатах: z[p(FA)] и z [p(H)]. Поскольку вероятности p(H) и p(FA) оценивались по частотам (т.е. мы имеем лишь их приблизительные значения), то точки, соответствующие z-преобразованиям, будут отклоняться от теоретической прямой (с наклоном 45 градусов) даже в том случае, если проверяемые предположения верны. Следовательно, надо провести прямую наилучшего приближения и проверить с помощью стандартных статистических средств, значимо или не значимо ее наклон отличается от 45°. Если отличие не значимо, исходные предположения могут считаться верными, а величина свободного члена в формуле прямой дает нам статистическую оценку d'. Разумеется, всем этим выводам должна предшествовать проверка того, является ли расположение экспериментальных точек хорошим приближением к прямой линии, т.е. необходимо провести статистический тест на линейность.

Допустим теперь, что удалось показать, что z-преобразованная РХ не является прямой с наклоном в 45 градусов. Тогда мы можем обратиться к более общему варианту нашей теоретической схемы: допустить, что σ_S распределения f(z/S) произвольна, но оба распределения нормальны. Очевидно, формула (14) сохраняет свою силу, так как *C* определяется только по p(FA). Изменения по отношению к случаю с $\sigma_{s,n} = 1$ появляются лишь в том месте, где распределение f(z/S) вместе с критерием *C* сдвигается влево до совмещения центра с нулевой точкой. Теперь мы уже не можем написать формулы (15) и (16), так как сдвинутое распределение описывается формулой: $\frac{1}{\sqrt{2\pi}} e^{-z^2/2\sigma^2}$. Однако, если мы вдобавок к сдвигу сожмем ось Z ровно в σ раз, то распределение приобретет нужную нам табличную форму. При этом критерий *C*, который после сдвига занял позицию *C* - *a* (мы уже не напишем d' вместо *a*), займет позицию $\frac{C-a}{\sigma}$. Итак:

$$p(H) = \frac{1}{\sqrt{2\pi}} \int_{\frac{C-a}{\sigma}}^{\infty} e^{-z^2/2} dx \quad \text{и} \qquad (18)$$

$$\frac{C-a}{\sigma} = -z[p[(H)]^{\frac{C-a}{\sigma}}. \qquad (19)$$

Сопоставляя (14) и (19) имеем:

$$z[p(H)] = \frac{z[p(FA)]}{\sigma} + \frac{a}{\sigma} \qquad (20)$$

Итак, если оба распределения нормальны, то график РХ в двойных нормальных координатах должен быть прямой линией с наклоном 1/σ (см. рис.11). Для проверки предположения о нормальности нужно оценить возможность описания экспериментальных точек линейной функцией или, (другими словами) “хорошесть” подгонки прямой линии к экспериментальным точкам.

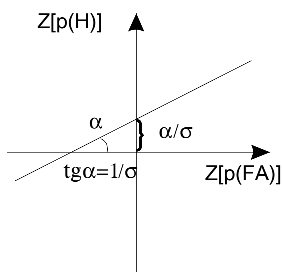


Рис.11. РХ в двойных нормальных координатах, $\sigma_S \neq \sigma_N$.

На основании статистических оценок предположение о нормальности отвергается, если даже наилучшая (в смысле метода наименьших квадратов, например) прямая плохо подходит к данным.

Предположим, что распределения f(z/S) и f(z/N) имеют одинаковые дисперсии, то есть РХ в двойных нормальных координатах является прямой линией с наклоном 1.

Положение каждой отдельной точки на РХ соответствует некоторому положению критерия *C*.

Можно показать, что при сделанных нами допущениях о нормальности распределений и равенстве дисперсий каждому положению C взаимно однозначно соответствует так называемое *отношение правдоподобия* (в точке *C*) – β, которое определяется как:

$$\beta = \frac{f(C / S)}{f(C / N)}\,. \qquad (21)$$

Здесь f(C/S) и f(C/N) представляют собой значения функций плотности вероятности f(X/S) и f(X/N), взятые в критической точке *C*. Отношение правдоподобия β характеризует то, во сколько раз правдоподобнее, что сенсорная репрезентация, равная по величине значению *C*, будет вызвана значащим стимулом, чем стимулом пустым.

По некоторым теоретическим соображениям положение критерия принято характеризовать именно этим значением β, а не самой величиной *C*.

Значения f(C/S) и f(C/N) легко найти, зная p(H) и p(FA). Для этого необходимо воспользоваться таблицей плотности нормального распределения: найти значения плотностей, соответствующие Z[p(H)] и Z[p(FA)] (что мы уже умеем делать). Эти значения обозначаются через f[p(H)] и f[p(FA)]. Таким образом:

$$\beta = \frac{f[p(H)]}{f[p(FA)]}\,. \qquad (22)$$

Оказывается, однако, что не обязательно искать f-преобразования для того, чтобы вычислить β. Вместо этого проще (и полезнее) вычислить lnβ прямо по z-преобразованным вероятностям. Дело в том, что в формулы, выражающие p(H) и p(FA) через d' и β, последняя входит только в форме lnβ (попытайтесь сами вывести эти соотношения):

$$z[p(FA)] = -\frac{\ln\beta + \frac{d'^2}{2}}{d'} \quad , \qquad (23)$$

$$z[p(H)] = -\frac{\ln\beta - \frac{d'^2}{2}}{d'} \, . \qquad (24)$$

Отсюда легко вывести формулу для вычисления lnβ:

$$\ln\beta = -\frac{z[p(H)] + z[p(FA)]}{2} \cdot d' \, . \qquad (25)$$

§ 3. Метод двухальтернативного вынужденного выбора (2АВВ)

В методе 2АВВ предъявления всегда осуществляются *парами*, причем предъявления в одной паре либо следуют друг за другом во времени, либо осуществляются одновременно, но ясно разделены пространственно. Одна пара всегда состоит из <S> и <N>, и это испытуемому известно, но какое именно из предъявлений (первое или второе, правое или левое и т.п.) содержит сигнал, а какое является пустым, должен определить испытуемый. Например, предъявляется пара линий, одна из которых наклонена, а другая вертикальна. Линии располагаются слева и справа от фиксационной точки и после каждого предъявления испытуемый должен решить, какая линия (слева или справа) имела наклон. Другой пример. Испытуемый слышит постоянный белый шум. Во время прослушивания дважды (скажем, с интервалом в полсекунды) загорается и гаснет (в течении 50 мс) индикатор начала и конца предъявления. В одном из двух предъявлений к шуму добавляется слабый тон частотой 1000 Гц, и задача испытуемого состоит в том, чтобы указать, в первом или во втором предъявлении присутствовала тональная добавка.

Чтобы различать варианты организации пары стимулов, условимся один из элементов пары называть “первым” и записывать на первом месте, а другой — “вторым” и записывать на втором месте. Таким образом пара может иметь либо форму <S,N>, либо форму <N,S>. Допустим, если в нашем первом примере наклонная линия находится слева, мы имеем <H,B>, а если справа — <B,H>, где В означает “вертикальна”, Н — “наклонна”. Соответственно, если испытуемый считает, что наклонная линия находится слева, то его ответ может быть записан как “<H,B>”. В общем случае матрица стимулов-ответов представима в форме:

	«Да, Нет»	«Нет, Да»
<S,N>	правильный ответ 1	ошибка 1
<S,N>	ошибка 2	правильный ответ 2

Во всех остальных отношениях 2АВВ ничем не отличается от метода “Да-Нет”. Если условиться идентифицировать пару по ее первому элементу, то можно даже не менять обозначений. Например,

$$P(S) = P(<S,N>), \qquad P(N) = P(<N,S>) = 1 - P(S).$$

Правильный ответ 1 можно условно считать попаданием и обозначать его условную вероятность через p(H)=p("Да","Нет"/<S,N>); ошибку 2 можно условно считать ложной тревогой и использовать обозначение p(FA)=p ("Да","Нет"/<N,S>) и т.д. Аналогично методу “Да-Нет” вводятся платежные матрицы, обратная связь, предварительная информация. Укажем, однако, на одно существенное отличие. Если в методе “Да-Нет” P(S) и платежная матрица таковы, что мы допускаем, что субъективные цены обеих ошибок (FA и O) одинаковы, то вовсе не необходимо, чтобы условные вероятности этих ошибок были равны. Или, что то же самое, нет оснований, вообще говоря, ожидать, что p(H) = p(CR). В методе 2АВВ, однако, пары <S,N> и <N,S> симметричны и при сделанных предположениях условные вероятности правильных ответов 1 и 2 должны быть равны. Это интуитивное соображение подкрепляется теоретической моделью, к изложению которой мы переходим. Но прежде введем новое обозначение. Условимся через p(C) (от английского correct — правильный) обозначать суммарную вероятность правильного ответа:

$$p(C) = P(S)\cdot p(H) + P(N)\cdot p(CR). \qquad (26)$$

Результаты 2АВВ называются *несмещенными*, если $p(H) = p(CR)$ или, что то же самое, $p(H)+p(FA)=1$.

Теоретическая модель 2АВВ является простым распространением модели, изложенной в предыдущем разделе. Мы сразу предположим, что все сделанные там допущения и упрощающие предположения сохраняют свою силу по отношению к <S> и <N> по отдельности, а когда <S> и <N> объединяются в пару, их сенсорные репрезентации независимы друг от друга, причем испытуемый никогда не путает, какому (“первому” или “второму”) члену пары соответствует данный образ. Каждый образ оценивается по интенсивности некоторого выбранного качества, так что образ пары оценивается по паре интенсивности сенсорного качества <X1,X2>, записанных в той же последовательности, что и стимулы. Если предъявляется <S,N>, то X1 имеет распределение f(X/S), X2 — распределение f(X/N). Если предъявляется <N,S>, то наоборот X1 распределяется по f(X/N), а X2 — по f(X/S). Имея <X1,X2>, испытуемый должен решить, первая или вторая интенсивность соответствует <S>. Естественным правилом решения здесь является следующее: берется разность X1-X2 и сравнивается с критическим значением C^*. Если X1- X2 > C^*, то дается ответ “Да, Нет”, если же X1- X2 < C^*, то “Нет, Да”. Как видим, C^* играет здесь ту же роль, что и критерий C в методе “Да-Нет”. Заметим, что разность берется всегда в одном и том же направлении, скажем от “первой” интенсивности ко “второй”, X1-X2, независимо от того, было ли предъявлено <S,N> или <N,S>. Начнем с рассмотрения случая предъявления <S,N>. Поскольку X1 и X2 суть случайные величины, то их разность тоже является случайной величиной, распределение которой мы обозначим через f(Δx/<S,N>). f(Δx/

<S,N>) есть плотность вероятности того, что X1 - X2 = Δx при предъявлении <S,N>. Эта функция однозначно определяется, если известны два распределения f(X/S) и f(X/N). Пусть теперь предъявлена пара <N,S>. Очевидно, что в этом случае разность X2 - X1 распределена точно так же, как разность X1 - X2 в первом случае, т.е. плотность вероятности события X2 - X1 = Δx/<N,S> равна плотности вероятности события X1 - X2 = Δx/<S,N>; но ведь событие X1 - X2 = Δx/<S,N> равносильно событию X2 - X1 = Δx/<N,S>. Мы получаем важное соотношение:

$$f(\Delta x/\langle S,N\rangle) = f(-\Delta x/\langle N,S\rangle), \tag{27}$$

где разность всегда берется от “первой” интенсивности ко “второй”, X1-X2. Соотношение (27) означает, что функции распределения f(Δx/<S,N>) и f(Δx/<N,S>) являются зеркально симметричными. В этом существенное отличие теоретической схемы для 2АВВ от теоретической схемы для метода “Да-Нет”: f(X/S) и f(X/N) могут быть сколь угодно непохожими друг на друга, но f(Δx/<S,N>) и f(Δx/<N,S>) являются зеркальными копиями. Введем в теоретическое представление критерий *C**. На рис. 12 заштрихованные области равны по площади вероятностям p(CR) и p(H). Легко видеть, что несмещенный 2АВВ, при котором *p(CR) = p(H),* будет иметь место только в случае *C* = 0.*

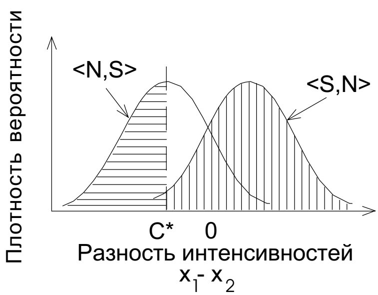


Рис.12. Геометрическая модель обнаружения стимулов в методе 2АВВ:
вертикальная штриховка - p(H); горизонтальная - p(CR);
C* - положение критерия принятия решения.

При отрицательных *C** испытуемый будет более часто правильно указывать сигнал, если сигнальное предъявление было “первым”, чем если оно было “вторым” (при этом говорят, что наблюдатель имеет предрасположение к “первому” стимулу). При *C*>0* испытуемый имеет предрасположение ко “второму” стимулу: *p(CR) > p(H)*. Двигая *C** справа налево и фиксируя различные пары *p(H), p(FA) (p(FA) = 1 - p(CR))*, мы можем построить кривую РХ для 2АВВ (рис. 13).

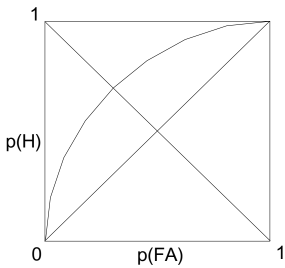


Рис.13. РХ для эксперимента по методу 2АВВ

В силу зеркальной симметричности распределений кривая РХ для 2АВВ всегда симметрична относительно побочной диагонали. Это следствие в принципе позволяет экспериментально проверить валидность схемы с оценкой разностей X1 - X2, но, к сожалению, строгое статистическое доказательство симметричности РХ провести довольно сложно. В эксперименте различные точки РХ можно получить, задавая асимметричные платежные матрицы (например, штрафуя за пропуск “первого” сигнала значительно больше, чем за пропуск “второго”), подавая одну комбинацию (например, <S,N>) чаще, чем другую и т.д. — совершенно аналогично методу “Да-Нет”.

До сих пор мы не использовали предположения о возможности монотонной трансформации X в Z, при которой f(X/S) и f(X/N) переходят в нормальные распределения f(Z/N) и f(Z/S). Если теперь это предположение принять и использовать разности Z1 - Z2, то можно показать следующее: если f(Z/N) имеет центр равным 0 и дисперисию равной 1, а f(Z/S) - центр в точке *a* и дисперсию равной σ , то f(ΔZ/<S,N>) и f(ΔZ/<N,S>) являются тоже нормальными распределениями с одной и той же дисперсией, равной $\sqrt{\sigma^2+1}$ и с центрами, соответственно, в точках *a* и -*a* (см. рис. 14).

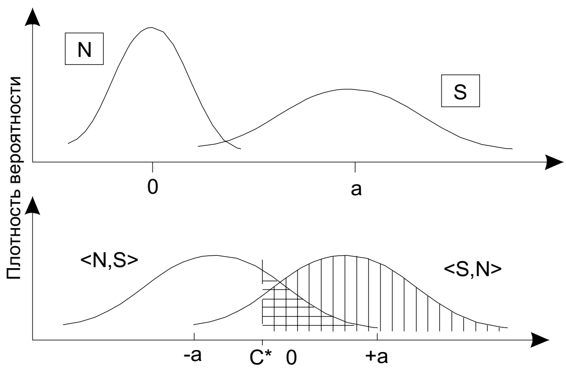


Рис. 14. Переход от распределений сенсорных эффектов, возникающих под действием пустого и значащего стимулов (<N> и <S>), к паре равновариативных распределений разности этих же сенсорных эффектов — <N,S> и <S,N>:

ось абсцисс — интенсивность сенсорного эффекта (верхний график) или разница интенсивностей сенсорных эффектов (нижний график); ось ординат — плотность вероятности соответствующих сенсорных эффектов

Рассмотрим, каковы соотношения между вероятностями p(H) и p(FA) при произвольном значении *C**. Для этого сдвинем левое распределение вместе с критерием до совмещения его центра с нулем и сожмем ось Z ровно в $\sqrt{\sigma^2+1}$ раз. Распределение после этого станет табличным, а критерий займет позицию $\frac{C^* + a}{\sqrt{\sigma^2 + 1}}$. Отсюда:

$$p(FA) = \frac{1}{\sqrt{2\pi}} \int\limits_{\frac{c^*+a}{\sqrt{\sigma^2+1}}}^{\infty} e^{-z^{2/2}} dz \quad \text{и} \qquad (28)$$

$$\frac{C^* + a}{\sqrt{\sigma^2 + 1}} = -z[p(FA)]. \qquad (29)$$

Вернемся теперь к исходной картинке и, сдвинув правое распределение вместе с критерием влево на *a* и, сжав z-ось в $\sqrt{\sigma^2 + 1}$ раз, получим:

$$p(H) = \frac{1}{\sqrt{2\pi}} \int\limits_{\frac{c^*-a}{\sqrt{\sigma^2+1}}}^{\infty} e^{-z^{2/2}} dz \quad \text{и} \qquad (30)$$

$$\frac{C^* - a}{\sqrt{\sigma^2 + 1}} = -z[p(H)]. \qquad (31)$$

Откуда:

$$z[p(H)] = z[p(FA)] + \frac{2a}{\sqrt{\sigma^2 + 1}}. \qquad (32)$$

Итак, в двойных нормальных координатах РХ для 2АВВ описывается прямой линией с наклоном 45 градусов (заметьте, при любой величине σ). Отсюда следует способ экспериментальной проверки предположения о нормальности f(z/S) и f(z/N) в методе 2АВВ: по z-преобразованным точкам РХ строится прямая наилучшего приближения, проверяется удовлетворительность приближения и незначимость отличия наклона от 45 градусов. Если дополнительно предположить, что σ = 1 , т.е. f(z/S) и f(z/N) имеют одинаковые дисперсии, то свободный член в формуле (32) станет равен $a\sqrt{2}$ (или, применяя стандартное обозначение, $d'\sqrt{2}$). В этом случае для разности z[p(H)] - z[p(FA)] в 2АВВ тоже иногда используют обозначение d' и пишут:

$$d'2ABB = \sqrt{2}d' \text{ "Да – Нет"} \qquad (33)$$

Часто это соотношение (не очень корректно) читается так: чувствительность в 2АВВ в $\sqrt{2}$ выше, чем в "Да-Нет". Этот вывод вряд ли покажется неожиданным для психолога, поскольку почти очевидно, что в условиях, где у испытуемого имеется возможность *сравнения*, результаты будут выше, чем в тех условиях, где такая возможность отсутствует (метод "Да-Нет").

В заключение мы остановимся на одном удивительном соотношении между 2АВВ и методом "Да-Нет". Мы знаем, что чувствительность (отличимость сигнального стимула от пустого) может быть измерена числом d' , если на распределении f(X/S) и f(X/N) наложено весьма жестко требование о существовании монотонной трансформации X®Z, переводящей эти распределения в два нормальных с равными дисперсиями. Если это требование не выполняется, но f(X/S) и f(X/N) могут быть переведены путем монотонной трансформации в два нормальных распределения с разными дисперсиями, то в методе "Да-Нет" чувствительность характеризуется уже парой чисел (*a,* s), что весьма неудобно, поскольку к парам чисел неприложимы оценки "больше-меньше", "возрастает-убывает" и т.д. Разумеется, в этом случае можно предложить какую-либо другую скалярную, (т.е. выразимую одним действительным числом) меру чувствительности (на рис. 15 показана одна такая мера, называемая d_{yn}), которая с формальной точки зрения будет являться скалярной функцией от *a* и s (например, $d_{YN} = \dfrac{a}{\sqrt{\sigma^2 + 1}}$).

Или можно обратиться к 2АВВ, взяв за меру чувствительности свободный член уравнения (32). Однако часто возникает вопрос, что делать в том случае, когда проверка отвергает предположение о нормальности? Существует ли какая-либо простая скалярная мера чувствительности, применимая при любых f(X/S) и f(X/N)? Такая мера действительно существует: *площадь под кривой РХ*. Интуитивно эта мера представляется весьма удачной. Она универсальна (применима к любой РХ) и всегда позволяет сказать, в каком сигнальном стимуле, S1 или S2, сигнал более обнаруживаем (в сопоставлении с одним и тем же N). Но у этой меры (обозначим ее U, см. рис. 16) есть существенный недостаток — для ее вычисления необходимо знать достаточно много точек РХ.

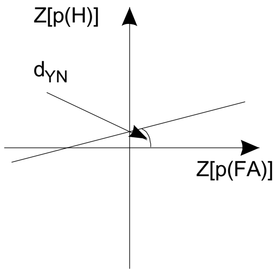


Рис. 15. Графическое представление меры чувствительности d_{YN} на РХ в двойных нормальных координатах

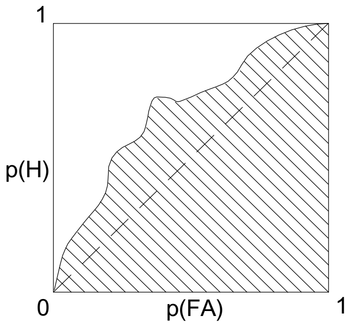


Рис. 16. Графическое представление меры чувствительности U на РХ в двойных нормальных координатах

Допустим, однако, что для некоторой пары <N> и <S> было проведено подробное исследование и вычислена мера U. Пусть теперь мы используем те же <S> и <N> в методе 2АВВ. Мы провели всего один эксперимент и получили (с точностью до статистических вариаций) следующий результат:

	«Да, Нет»	«Нет, Да»
<S,N>	p	1 - p
<N,S>	1 - p	p

Результаты показывают, что выбор является несмещенным: p(H) = p(CR). Мы знаем, что в этом случае общая вероятность правильного ответа P(C) (см. формулу (26)) равна p. Удивительное соотношение между “Да-Нет” и 2АВВ, о котором идет речь, состоит в том, что если изложенная модель обнаружения верна, то должно быть U = p. Другими словами: в несмещенном случае $P(C)_{2ABB} = = U_{\text{"Да-Нет"}}$. Таким образом, в качестве хорошей и простой (пожалуй, самой простой) меры чувствительности в 2АВВ может использоваться процент правильных ответов P(C).

§ 4. Метод оценки

Этот метод может быть использован как модификация метода “Да-Нет” и как модификация метода 2АВВ. Здесь будет изложен только первый вариант, поскольку перенесение его на случай 2АВВ является тривиальным.

Как мы уже знаем, в ряде случаев (для проверки гипотез о форме распределений или для вычисления таких мер чувствительности, как U) требуется РХ по достаточно большому количеству точек. Для получения нескольких точек РХ методом “Да-Нет” необходимо несколько раз провести эксперимент с одной и той же парой <S> и <N>, но с различными параметрами организации эксперимента, такими как P(S), платежная матрица и т.п. Каждый эксперимент должен содержать большое количество предъявлений для того, чтобы, во-первых, можно было исключить первые пробы, в которых схема соответствия еще не установилась, и, во-вторых, чтобы частоты событий (“Да”/S) и (“Да”/N), высчитанные по оставшимся пробам (асимптотический уровень), достаточно точно соответствовали вероятностям p(H) и p(FA). Более того, поскольку от эксперимента к эксперименту чувствительность наблюдателя к данному сигналу может меняться, эксперимент с одними и теми же параметрами организации желательно повторить несколько раз на разных этапах (скажем, ближе к началу, середине и концу) всей серии экспериментов. Все это довольно громоздкая работа. Метод оценки (МО) дает нам возможность получить несколько точек РХ в результате только одного эксперимента, хотя его объем, обыкновенно, превышает объем одного эксперимента “Да-Нет”.

Процедура метода оценки (МО) отличается от метода “Да-Нет” только тем, что после каждого предъявления вместо ответа “Да” или “Нет” испытуемый указывает *степень его уверенности* в наличии/отсутствии сигнала в этом предъявлении. Например, “совершенно уверен, что сигнал был”, “уверен, что сигнал был”, “скорее был, чем не был”, “не могу выбрать”, “скорее не был, чем был”, “уверен, что сигнала не было”, “совершенно уверен, что сигнала не было”. Эти 7 категорий естественно обозначить числами в том же порядке: 3, 2, 1, 0, -1, -2, -3. В методе оценки уверенности набор категорий всегда задается испытуемому заранее и обычно кодируется некоторой числовой системой. Иногда используется процентная шкала, когда испытуемый говорит о сигнале: “На 50% был”, “На 100% был” (точно был), “На 10% был”, “На 0% был” (точно не был). В этом случае либо испытуемого просят пользоваться только определенными (например, только круглыми: 0, 10, 20 ...%) числами, либо он может называть произвольные проценты (скажем, 78%), но потом ответы объединяются в несколько групп (например, все числа меньше 5% — в группу 0, все числа между 5 и 15 — в группу 10% и т.д.). Для конкретности предположим, что испытуемому заданы 7 категорий, названных в нашем примере. Обыкновенно эксперимент проводится без платежной матрицы или с симметричной платежной матрицей и с P(S) = P(N) = 0.5. Результаты эксперимента могут быть представлены в виде следующей таблицы (см. табл. 5).

Таблица 5

Теоретические результаты эксперимента с использованием метода оценки

Области	-3	-2	-1	0	1	2	3
S	p(-3)	p(-2)	p(-1)	p(0)	p(1)	p(2)	p(3)
N	q(-3)	q(-2)	q(-1)	q(0)	q(1)	q(2)	q(3)

P(n), n=-3,...+3, есть оценка условной вероятности P(n/S), получаемая путем деления числа всех случаев, когда предъявлялось <S> и был дан ответ “n”, на число всех предъявлений <S>. Аналогично q(n) есть оценка условной вероятности P(n/N). Теоретическое осмысление этих данных в рамках модели, изложенной в двух предыдущих разделах, состоит в предположении, что если испытуемому заданы *K* категорий (от полной уверенности в отсутствии до полной уверенности в наличии S), то он так же, как и в условиях эксперимента “Да-Нет”, базируется на интенсивности некоторого сенсорного качества, но делит ее не на две, а на *K* областей, как показано на рис. 17.

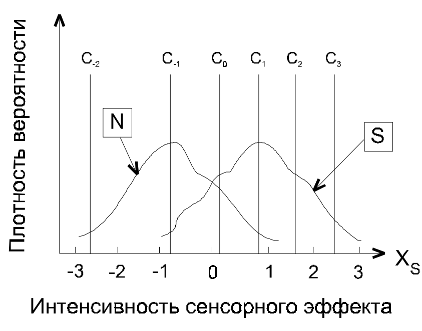


Рис. 17. Модельное представление ситуации обнаружения сигнала в методе МО

Как видим, совсем необязательно, чтобы границы между областями разных ответов следовали через равные интервалы или каким-нибудь закономерным образом: единственное, что предполагается — что область ответа R_1 лежит левее области ответа R_2, если $C_1 < C_2$. Итак, если выбранное качество сенсорного образа имеет интенсивность, лежащую между C_0 и C_1, то испытуемый дает ответ “0”, если интенсивность лежит правее C_3 — то “3” и т.д.

Теперь приведем следующее рассуждение. Допустим, что те же стимулы <N> и <S> используются в эксперименте “Да-Нет”, причем критерий C будет последовательно помещаться в позиции C_3, C_2, C_1, C_0, C_{-1}, C_{-2}. При каждом положении критерия будем вычислять соответствующую пару p(H) и p(FA). Вероятность p(H) равна площади под кривой *f(X/S)*, лежащей правее C, а p(FA) равна площади под кривой *f(X/N)*, лежащей правее C. Обозначим площадь под кривой *f(x/S)* между C_i и C_{i+1} (i = -2, -1 ... 2 в нашем случае на рис. 17) через A_S (C_i, C_{i+1}), а площадь, лежащую правее C_i - через A_S (C_i, C_∞). Для кривой *f(X/N)* — аналогичные обозначения: A_N (C_i,C_{i+1}) и A_N (C_i, C_∞). Если критерий C помещен в позицию C_i , то p(H) = A_S (C_i, C_∞), p(FA) = A_N(C_i, C_∞). С другой стороны, ясно, что p(i) - вероятность ответа «i» при предъявлении S, равна A_S (C_i, C_{i+1}), если i < 3 и равна A_S (C_3, C_∞), если i = 3. Аналогично q_i = A_N (C_i,C_{i+1}), если i < 3 и A_N (C_3, C_∞), если если i = 3. Но, очевидно, что, $A_S(C_0, C_\infty) = = A_S(C_0, C_1) + A_S(C_1, C_2) + A_S(C_2, C_3) + A_S(C_3, C_\infty)$, и аналогично раскладываются любые другие A_S (C_i, C_∞) и A_N (C_i , C_∞).
Следовательно, мы получаем следующую цепочку равенств (табл. 6):

Таблица 6

Способ расчета p(H) и p(FA) по полученным данным в методе МО

Положение критерия	p(H)	p(FA)
C_3	p(3)	q(3)
C_2	p(2) + p(3)	q(2) + q(3)
C_1	p(1) + p(2) + p(3)	q(1) + q(2) + q(3)
C_0	p(0) + p(1) + p(2) + p(3)	q(0) + q(1) + q(2) + q(3)
C_{-1}	p(-1) + p(0) + p(1) + p(2) + p(3)	q(-1) + q(0) + q(1) + q(2) + q(3)
C_{-2}	p(-2) + p(-1) + p(0) + p(1) + p(2) + p(3)	q(-2) + q(-1) + q(0) + q(1) + q(2) + q(3)

Теперь мы имеем 6 пар вычисленных p(FA) и p(H) и, следовательно, имеем 6 точек РХ. Взяв больше категорий, мы построим РХ более подробно, но слишком большое число категорий требует очень длительного эксперимента (надо, чтобы каждая категория встречалась не слишком редко, иначе частота будет плохой оценкой вероятности) и поэтому на практике встречается не часто.

Методические рекомендации по выполнению учебных заданий по теме “Методы обнаружения сигнала”

На первом занятии, которое проходит в форме семинара, проводится обсуждение основ психофизической теории обнаружения сигнала (ТОС), являющейся рабочим инструментом современной психофизики. К этому занятию студенту необходимо прочесть данную главу учебного пособия. В качестве альтернативной и/или дополнительной литературы может быть рекомендована глава 7 книги К.В Бардина (1976). Для студентов, имеющих более солидную математическую подготовку и дополнительный интерес к освоению методов обнаружения сигнала, будут полезны 1—3 главы монографии Дж. Игана(1983). Часть первого и второе занятия посвящаются планированию предстоящего эксперимента, освоению программного обеспечения, с помощью которого проводится отработка учебного задания (см. *Приложение 2*), и выполнению тренировочных серий эксперимента. Третье (и при необходимости четвертое) занятие отводится для проведения основных серий эксперимента и подготовки отчета.

Предполагается, что студент уже имеет элементарные навыки самостоятельной работы на IBM-совместимом персональном компьютере.

Основное внимание при обсуждении теоретических основ ТОС необходимо обратить на те теоретические предположения, которые делаются в психофизической теории обнаружения сигнала, на отличие данного подхода к измерению чувствительности от классического фехнеровского подхода. Известную трудность при изложении данной модели обнаружения сигнала составляют особенности ее формально-математического описания, тем не менее они не выходят за рамки тех минимальных знаний об интегральном и дифференциальном исчислении, которые были получены студентами на 1-м курсе. Кроме того, в ходе освоения материала нетрудно отделить собственно психологические предположения и ограничения, накладываемые моделью в силу упрощения или даже примитивизации описываемой реальности, и следующие из этого математические допущения. Нужно себе четко представлять, что попытка формально-математического описания даже таких “низкоуровневых” процессов как обнаружение или различение простых сенсорных сигналов, сталкивается с необходимостью “вынести за скобки”, нивелировать большинство таких детерминант сенсорно-перцептивного процесса, как колебания внимания, когнитивно-стилевые особенности человека, индивидуальность его мотивации и др. Хорошо это или плохо, но большинство попыток модельного описания психических процессов, представленных в современной литературе, в той или иной степени приводит к аналогичному результату (см., например, модели памяти Аткинсона или когнитивные варианты современных моделей мотивации, где делаются более глобальные и далеко идущие предположения и ограничения в описании куда более сложной моделируемой реальности).

При проработке материала следует обратить внимание на двухэтапность описываемого процесса обнаружения сигнала. Первый этап связан непосредственно с *сенсорной репрезентацией* действующих стимулов, т.е. с отражением стимульной энергии в величину вызванного ими ощущения; и как результат — постулируемое распределение (на оси X) интенсивности сенсорных эффектов или, что тоже самое, — ощущений заданного в инструкции сенсорного качества. Основные детерминанты этого (сенсорного) этапа – физические характеристики стимуляции и особенности анализаторной системы. Сразу же отметим, что делаемое допущение о *нормальности* гипотетического распределения на сенсорной оси есть не только дань простоте при математическом моделировании, но и следствие обобщения опыта многочисленных пороговых измерений, известного в истории психологии как “фи-гамма” гипотеза. В этой связи полезно вспомнить, почему данную модель считают “непороговой”. Такое определение основывается на принятии за основу *вероятностного* принципа отображения энергии стимула в величину ощущения (сравните с детерминистическим определением порога как границы в классической психофизике), из чего следует отсутствие как такового порога на сенсорной оси и, следовательно, непороговый принцип работы сенсорной системы.

Второй этап характеризует процесс принятия решения о полученном ощущении и связан с *внесенсорной детерминацией* процесса обнаружения (различения) сигнала. *Критерий принятия решения* является тем интегральным показателем, который и определяет окончательный результат процесса обнаружения сигнала. Обычно при описании данной модели критерий наблюдателя помещают на сенсорной оси, тем самым указывая на его природу. Подчеркнем, что, являясь по своей сути сенсорным эталоном обнаруживаемого сигнала, стандартом для сравнения с текущим стимулом, критерий формируется не только и не столько под действием стимуляции (например, в ходе тренировки), но во многом зависит от несенсорных факторов. Различного рода экспериментальные установки и ожидания, сформированные инструкцией экспериментатора и/или самоинструкцией, влияют на выбор стратегии испытуемого при принятии решения о наличии сигнала в очередной пробе. В *Приложении 1* приведены дополнительные сведения о различных критериях оптимальности принятия решения, используемых в современной психофизике, и описан принятый в ТОС критерий наблюдателя, основанный на оценке отношения правдоподобия. Расчет отношения правдоподобия — один из основных способов параметрического (т.е. основанного на законах постулируемого в ТОС нормального распределения сенсорных эффектов) измерения критерия наблюдателя. Следует особо подчеркнуть, что сам математический аппарат, описывающий работу человека (или кибернетического устройства) с различными критериями, пришел в психологию из математической теории игр, и является не более, чем формальным описанием тех гипотетических процессов принятия решения, которые имеют место в ситуациях повышенной неопределенности. Очевидно, что задача обнаружения порогового сигнала, когда наблюдатель работает на пределе своих сенсорных способностей, представляет собой такую ситуацию. Учитывая формальный характер описания работы наблюдателя с определенным критерием, апелляцию к определенному критерию (например, критерию по типу отношения правдоподобия) нужно рассматривать не более, чем формализованное (модельное) описание результата некоторых гипотетических процессов принятия решения. В этом смысле психологический анализ деятельности наблюдателя должен идти от содержательной психологической интерпретации использования им того или иного критерия, а не от вычисления определенной математической функции, описывающей критерий, которая сама по себе может быть свободна от психологического содержания.

Задание 1. Обнаружение зрительного сигнала методом «Да-Нет»

Цели задания. *1. Практическое освоение метода «Да-Нет» на примере обнаружения зрительного сигнала. 2. Исследование динамики d' и* β *в зависимости от влияния несенсорных факторов.*

Методические замечания по планированию и проведению эксперимента.

При планировании предстоящего исследования стоит обратить особое внимание на важность *тренировочных серий* эксперимента и вспомнить, каким требованиям должен удовлетворять идеальный испытуемый (наблюдатель). Прежде всего еще раз подчеркнем, что в предлагаемой модели описывается ситуация обнаружения сигнала порогового уровня, следовательно в ходе тренировочных серий необходимо подобрать соответствующие параметры обнаруживаемого сигнала. В компьютерной программе стимуляции (см. Приложение 2) предлагаются на выбор различные сигнальные и несигнальные стимулы, например: обнаруживать букву R на фоне L, I на фоне 1 или Q среди O. Естественно, принимая во внимание индивидуальные особенности зрения испытуемого, следует подобрать такие стимулы, которые будут с трудом отличаться друг от друга, и в этом смысле, по-видимому, вариант R и L (это достаточно хорошо различимые конфигурации) будет адекватен лишь для тех студентов, у кого не очень хорошее зрение. В противном случае, как показывает наш опыт, даже при минимальном времени экспозиции стимулов на экране дисплея после хорошей тренировки некоторые испытуемые показывают практически 100%-е обнаружение такого сигнала. Интересно, что поначалу это может показаться весьма сомнительным, но поработав 15—20 минут, как правило, все убеждаются, что тренировка идет и, несмотря на невысокую уверенность каждого отдельного ответа в прошедшей серии, результат обнаружения почти 100%-й. И, следовательно, время предыдущих тренировочных серий потрачено не оптимально. Таким образом, с самого начала нужно четко представлять себе, что следует выбрать такие стимулы и такую их длительность, чтобы обеспечить *пороговый уровень* обнаружения сигнала. Для более четкой ориентации введем операциональный критерий "пороговости" обнаружения сигнала: индекс сенсорной чувствительности d' должен быть в диапазоне от 1 до 2, что соответствует вероятности попаданий явно меньшей 1 и вероятности ложных тревог, превышающей 0. Например, если тренировочные серии проводятся при априорной вероятности предъявления сигнала, равной 0.5, то соответствующие значения вероятностей попаданий и ложных тревог будут приблизительно такими: p(H) - от 0.7 до 0.8, а p(FA) - от 0.1 до 0.3.

Следующий немаловажный момент касается вопроса о достижении испытуемым *асимптотического* (предельного) уровня обнаружения порогового сигнала, а именно, достиг ли он того предельного уровня тренировки, когда со временем практически не происходит существенных изменений d'. Самым простым подтверждением достижения асимптотического уровня обнаружения будет относительное постоянство показателей обнаружения в 3—4 следующих друг за другом тренировочных сериях при *неизменных* стимульных параметрах. Полезно также посмотреть, как изменяется среднее время реакции (ВР) и его вариативность. Стабилизация величины среднего ВР и его разброса служит хорошим доказательством выхода испытуемого на асимптотический уровень обнаружения. В табл. 7 приведены реальные результаты тренировочной серии эксперимента (данные студента Е.К., 1994 г.), показывающие достижение к шестой серии асимптотического уровня обнаружения сигнала.

Таблица 7

Результаты тренировочных серий (задача – обнаруживать Q на фоне O, длит. стимула – 250 мс, МСИ – 2000 мс)

Номер серии	P(H)	P(FA)	ВР, мс	d'
1	0.62	0.34	645	0.72
2	0.68	0.3	664	0.99
3	0.74	0.23	560	1.38
4	0.78	0.18	568	1.69
5	0.81	0.20	548	1.72
6	0.78	0.20	573	1.61

Естественным будет вопрос о пределах вариабельности индекса d'. Укажем, что строгая статистическая оценка различий d', полученных в разных сериях одного эксперимента или разных экспериментах производится с использованием критерия хи-квадрат (можно воспользоваться специальной программой hi_sq.exe, которая находится в той же директории, что и основная программа yes_no.exe), однако для быстрой оценки существенности полученных различий можно использовать чисто эмпирический критерий, проверенный на практике: *25—30%-е* различие индексов d', как правило, не значимо. Несмотря на то, что данная величина на первый взгляд кажется достаточно большой, следует учесть, что d' оценивается вероятностно и является производным показателем, зависящем как от P(H), так и от P(FA), которые, в свою очередь, представляют собой тоже случайные величины, оцениваемые в опыте также вероятностно. Таким образом, следует обратить особое внимание на достоверность оценки этих 2-х вероятностей, что непосредственно определяется ***количеством предъявляемых стимулов*** — сигнальных и несигнальных. Интуитивно ясно, что по 5—10 пробам невозможно оценить вероятность появления какого-либо события; можно показать, что по 85—100 (т.е общее число проб = 190—200 при P(S) = 0.5) предъявлениям сигнальных и шумовых проб оценка вероятности правильного обнаружения и ложной тревоги становится статистически надежной. Из данных соображений и следует исходить при решении вопроса об определении минимального количества проб в каждой серии. Естественно, следует учитывать и значение априорной вероятности появления сигнальной или шумовой проб: чем меньше выбирается вероятность данного стимула (сигнального либо шумового), тем большее количество проб в данной серии следует предъявить испытуемому. Поэтому даже в тренировочных сериях (кроме самых предварительных) не следует “экономить” на количестве проб. Использование малого количества проб в серии может привести к следующему результату: показатели обнаружения сигнала (P(H), P(FA) и как интегральный показатель — d']) могут сильно изменяться от серии к серии, а мы не сможем определить, в чем же причина этой вариабельности — либо в том, что имеет место тренировка, либо это просто случайные вариации оцениваемых вероятностей от серии к серии. Данное замечание следует учитывать особо в том случае, если в основном эксперименте в качестве несенсорного фактора варьируется априорная вероятность предъявления сигнальной пробы (в отличии от варианта с использованием платежной матрицы). Как показывает практика, при низких значениях вероятности (0.1 и 0.9) следует предъявлять не менее 450—500 проб, при вероятностях 0.2 и 0.8 − 300—350, при равновероятном предъявлении − 190—200.

Важное значение при выполнении данного задания имеет учет фактора *утомления*. Эксперимент проходит достаточно длительное время, поэтому после каждой серии необходимо устраивать небольшой *перерыв* для отдыха.

Особое внимание следует уделить планированию основного эксперимента. Основная цель данного учебного задания — провести модельный эксперимент в рамках ТОС и познакомиться с методом “Да-Нет”. Таким образом, непосредственная задача эксперимента заключается в построении РХП, т.е. в варьировании несенсорных факторов, задающих несколько различных критериев принятия решения. При выборе конкретного приема экспериментального воздействия (использование априорной вероятности, платежной матрицы либо инструкции) стоит учесть, что для неопытного (наивного) испытуемого большое значение имеет правильное представление о критерии оптимальности выполняемой задачи и однозначное понимание и принятие задачи эксперимента[1]. В этом смысле более предпочтительным оказывается использование различных *платежных матриц или варьирование априорной вероятности.* Эти приемы наиболее прямо и наглядно показывают испытуемому, как следует изменить стратегию обнаружения сигнала, чтобы оптимальным образом выполнить свою задачу — эффективнее обнаруживать сигнал в ситуации неопределенности. И в том, и в другом случае испытуемый должен четко и однозначно представлять себе, что влечет за собой определенное изменение априорной вероятности или платеж-

---

[1] Дополнительные сведения о различных критериях принятия решения и критериях оптимальности приведены в Приложении 1.

ной матрицы. Так, еще до начала основного эксперимента полезно прикинуть, как следует себя вести в сериях с различной априорной вероятностью появления сигнального стимула, и что же реально происходит, когда в одной серии P(S)= 0.1, а в другой P(S) меняется на 0.9. Очевидно, что изменение априорной вероятности формирует соответствующие изменения ожиданий испытуемого в отношении последовательности предъявляемых в данной серии стимулов, что немаловажно в ситуации повышенной неопределенности (т.е. далеко не 100%-й обнаружимости сигнала). Иначе говоря, когда вы не очень-то уверены, какой из 2-х сигналов был предъявлен, и у вас возникает сомнение, то важным несенсорным признаком стимуляции оказывается знание вероятности предъявления сигнального стимула, которое поможет правильно *угадать*. А теперь давайте прикинем, насколько оптимально следовать таким правилам "игры". Примем условно, что явно сомнительных ощущений из 200 проб оказалось 100, т.е. половина. Допустим, что в данной серии P(S)=0.9. Тогда становится ясно, что даже обычное гадание в этих "сомнительных" 100 пробах на основании простого учета вероятности появления сигнала (ведь шанс правильно угадать — 90 из 100 !) может принести наблюдателю заметную пользу и, что тоже не маловажно, снять излишнюю напряженность в работе (ведь гадаем-то на основании трезвого расчета). Несложно "проиграть" аналогичную ситуацию "со знаком минус" — когда P(S)=0.1, и распространить эту стратегию на другие значения априорной вероятности.

В том случае, когда студенты (экспериментатор и испытуемый составляют симметричную пару) выбирают в качестве экспериментального воздействия платежную матрицу, то ситуация становится еще более прозрачной — ведь каждый четко знает, сколько в данной серии стоит каждый тип ответов. Меняя цены наград и штрафов (как правило, оба партнера договариваются об этом сами, прикидывая максимально возможный выигрыш и проигрыш), не очень сложно построить 5—7 платежных матриц, градуально задающих строгость/либеральность критерия принятия решения об обнаружении сигнала. Так, сильно штрафуя ложные тревоги по отношению к пропускам сигнала и умеренно вознаграждая правильные ответы, однозначно поощряем строгий критерий. И наоборот, значительное поощрение правильных обнаружений с существенным наказанием пропусков и мягким наказанием за ложные тревоги объективно подталкивает испытуемого к использованию либерального критерия. Выбрав достаточно большой масштаб изменения наград и штрафов, не представляет особого труда составить ряд платежных матриц от явно строгого до явно либерального критерия. Стоит подчеркнуть, что в данном эксперименте партнеры должны строго соблюдать следующее правило: подсчитывать свои выигрыши (проигрыши) после каждой серии, сравнивать их, а разницу фиксировать в протоколе, чтобы было точно понятно, кто в данной серии выиграл и сколько. Опыт показывает, что целесообразно использовать реальные деньги, а не просто очки или баллы. Нужно помнить, что в реальном психофизическом эксперименте испытуемым всегда платят деньги, так что лучше не нарушать традицию. Конечно, стоит заранее договориться и ограничить максимально возможный размер проигрыша и выигрыша при неоптимальной и оптимальной стратегиях, соответственно.

И еще несколько слов по поводу планирования эксперимента. Стоит помнить о двух основных факторах, мешающих проведению нашего эксперимента и способных исказить его результат — это *тренировка и утомление.* Учет и того, и другого очень важен, поскольку эксперимент состоит из нескольких серий, распределенных во времени. Каким образом избежать возможного влияния этих факторов? Для этого используют прием, называемый *позиционным уравниванием.* Каждую серию эксперимента (допустим, что их будет 5 — по числу разных априорных вероятностей) разбивают на две подсерии и эти половинки располагают в эксперименте в следующем порядке:

P(0.1) - P(0.3) - P(0.5) - P(0.7) - P(0.9) - P(0.9) - - P(0.7) - P(0.5) - P(0.3) - (0.1). Задавая такой порядок следования отдельных серий эксперимента, мы тем самым уравниваем возможное влияние факторов тренировки и утомления на деятельность испытуемого, усредняя показатели обнаружения сигнала по двум соответствующим половинкам. Резон здесь такой: для первой половины каждой серии минимально утомление, но и тренировка минимальна тоже, для второй половины — наоборот. Поэтому, усредняя данные по двум сериям, мы тем самым уравниваем разнонаправленное влияние этих факторов на результаты обнаружения сигнала. Кроме того, усреднняя данные, взятые из разных временных срезов эксперимента, мы отчасти компенсируем влияние других неконтролируемых случайных факторов (внешние помехи, случайные колебания стимуляции и т.д.).

Оценивая возможное влияние различных нежелательных факторов на показатели обнаружения сигнала, сделаем еще несколько замечаний относительно проведения эксперимента. Во-первых, весь эксперимент следует проводить *на одном и том же компьютере*. Во-вторых, если весь эксперимент не получается провести *в один день*, то в следующий раз необходимо провести тренировочную серию и убедиться в том, что вы достигли прежнего уровня обнаружения сигнала. В-третьих, ни в коем случае *не меняйте параметры стимуляции* по ходу основного эксперимента, помня, что вы имеете дело только с изменением несенсорных факторов, будь то априорная вероятность или платежная матрица, в то время как детерминанты сенсорной части процесса обнаружения должны оставаться неизменными.

Обработка и интерпретация результатов.

По окончании каждой серии, студент получает файл с результатами обнаружения сигнала. Целесообразно записывать *в отдельный протокол* значения основных показателей обнаружения сигнала: P(H), P(FA), d', β, среднее ВР, а также параметры стимуляции (длительность стимула, количество стимулов в серии) и варьируемые несенсорные факторы — априорную вероятность или вид платежной матрицы. Кроме того, после каждой серии полезно делать хотя бы короткие записи *самоотчетов*, где фиксировать свои впечатления о прошедшей серии.

По итогам эксперимента необходимо рассчитать усредненные по двум половинам каждой серии вероятности попаданий и ложных тревог и построить *РХП в линейных и z-координатах.* Если в линейных координатах РХП имеет достаточно стандартный вид (сравните с рис. 8), то проведите через все точки "на глазок" плавную кривую. Имеет смысл построить для каждой точки РХП *гипотетический 10—20% доверительный интервал*, и проводить наилучшую кривую с учетом такого разброса оценок каждой вероятности (это не совсем корректно в смысле строгой статистики, но, тем не менее, позволит вам почувствовать проблему вероятностной подгонки полученных данных под ожидания модели). На графике в z-координатах следует нанести все экспериментальные точки и, следуя ожиданиям модели, провести через них прямую линию. При решении проблемы, как провести через все точки наилучшую прямую (для РХП в z-координатах), следует воспользоваться методами регрессионного анализа. Задача подгонки прямой линии под экспериментальные точки решается следующим образом (принимая во внимание, что и по оси абсцисс и по оси ординат мы имеем оценки функции, необходимо построить наилучшую прямую с учетом вероятного разброса оценок по *каждой* из них). Нужно построить линейную регрессию z(H) по z(FA) − это наилучшая прямая с учетом разброса по X, и аналогичную регрессию z(FA) по z(H) — это наилучшая прямая с учетом разброса по Y, и изобразить обе эти прямые в осях z(H) — z(FA). Проведя биссектрису угла между этими прямыми, мы получим наилучшую (с точки зрения метода наименьших квадратов) прямую с учетом разброса оценок как z(H), так и z(FA). Для решения этой задачи можно использовать статистический пакет "Stadia": введите в первую колонку z-оценки ложных тревог, а во вторую - попаданий; после этого выберете в меню статистических методов рубрику "Регрессионный анализ", а в ней опцию — простая регрессия (тренд). После входа в соответствующее меню нужно выбрать линейную модель и произвести два раза регрессионный анализ — z(H) по z(FA) и z(FA) по z(H) (не забудьте списать с экрана рассчитанные коэффициенты полученных линейных функций). Целесообразно также посмотреть полученные графики на экране компьютера. В том случае, если оба варианта подгонки статистически достоверно описываются линейными функциями (см. заключение «Stadia» внизу экрана результатов), то с большой долей вероятности можно считать, что РХП в двойных нормальных координатах имеет форму прямой[1]. Таким образом проверяется первое основное предположение модели *о нормальности* распределения сенсорных эффектов. Для проверки второго предположения *о равновариативности* сигнального и шумового распределений нужно оценить угол наклона *прямой* РХП. Исходя из опыта, можно принять, что хорошим соответствием ожидаемому наклону в 45 градусов будет разброс ± 5—7 градусов. Однако можно сделать такую проверку и более строго, для чего достаточно всего лишь оценить гипотезу о равенстве дисперсий оценок по обоим осям – z(H) и z(FA), ведь при равенстве дисперсий эта прямая очевидно пройдет под углом 45 градусов! Для этого можно воспользоваться статистическим критерием Фишера в меню описательной статистики системы «Stadia». В том случае, если расчеты показывают, что дисперсия значений переменной z(H) достоверно не отличается от дисперсии переменной z(FA), можно принять гипотезу о наклоне прямой в 45 градусов. В противном случае это предположение отвергается.

В обсуждении результатов эксперимента следует обратить особое внимание на то, как изменялись показатели сенсорной чувствительности (d') и критерия (β) в разных сериях опыта и сопоставить их динамику с предположениями ТОС. В случае заметных расхождений следует дать содержательную интерпретацию таким различиям (при этом имеет смысл обратиться к записям самоотчетов). В том случае, когда в одной-двух сериях

[1] Ту же самую гипотезу можно проверить, рассчитав в статистическом пакете коэффициент корреляции пяти пар значений z(H) и z(FA) и оценив его статистическую достоверность, т.е. отличие от нуля.

получены результаты, сильно отличающиеся от ожидаемых, целесообразно эти серии переделать.

Задание 2. Обнаружение тонального сигнала на фоне шума методами двухальтернативного вынужденного выбора и оценки

Цели задания. *1. Практическое освоение методов на примере обнаружения акустического сигнала. 2. Сопоставление разных методов и мер, предлагаемых для оценки сенсорной чувствительности.*

Методика

Аппаратура. Звуковые сигналы предъявляются испытуемому через аудиометрические головные телефоны (например, “ТД-6” или “TDH-39”). Синтез и предъявление звуковых стимулов осуществляется с помощью прецизионного генератора аудиометрических частот[1], управляемого персональным компьютером. Управление стимуляцией, сбор ответов испытуемого и оперативная обработка полученных данных осуществляются компьютерными программами *2abb.exe* и *cr.exe.*

Стимуляция. Звуковые сигналы представляют собой отрезки широкополостного белого шума, к части из которых “примешан” тональный сигнал частотой 1000 Гц. Длительность звуковой посылки — 100 мс, интенсивность — 70—80 дБ по международной шкале SPL (шкала уровней звукового давления, где нулевому уровню соответствует величина среднего абсолютного порога слышимости).

---

[1] Программируемый звуковой генератор должен обеспечивать возможность регулировки интенсивности тональных сигналов и белого шума с дискретностью не хуже, чем ± 0.05 дБ. Для организации эксперимента можно воспользоваться также и стандартной звуковой картой к персональному компьютеру (типа “Sound Blaster”), с помощью которого с эталонного генератора записываются необходимые звуки, которые в дальнейшем и предъявляются в качестве стимулов.

Интенсивность тональной добавки регулируется с дискретностью ± 0.1 дБ.

В эксперименте по методу 2АВВ в каждой пробе “сигнальный” и “шумовой” стимулы предъявляются парами, с интервалом 500 мс. В опыте по методу ОУ в каждой пробе предъявляется только один стимул (сигнальный или шумовой).

Перед каждой пробой на экране дисплея в качестве сигнала “Внимание” предъявляется порядковый номер пробы.

Процедура. Каждый студент участвует в эксперименте в качестве испытуемого. Группа студентов делится пополам. Одна гполовина руппы сначала делает серию 2АВВ, потом ОУ, другая половина группы — наоборот. В обоих опытах в сигнальной пробе используется одно и то же отношение сигнал/шум, найденное в тренировочной серии. Если весь эксперимент проводится в один день, то тренировочная серия проводится лишь перед первым опытом, а перед вторым можно ограничиться лишь небольшой серией (40—50 проб), чтобы познакомиться со стимульной парадигмой и четко понять инструкцию. Если эксперимент продолжается в другой день, то перед началом следующего опыта рекомендуется провести хотя бы небольшую тренировочную серию (около 100 проб). В том случае, когда между двумя опытами прошел достаточно большой промежуток времени, стоит подумать о более длительной тренировочной серии, чтобы убедиться в достижении прежнего уровня продуктивности обнаружения сигнала.

1. Метод вынужденного выбора. Процедура опыта.

Опыт состоит из тренировочной и основной серий. В тренировочной серии испытуемый знакомится со стимульными условиями и процедурой эксперимента. В первой (ознакомительной) ее части предъявляются 20 проб (10 сигнальных и 10 несигнальных) с высоким *отношением сигнал/шум* в сигнальной пробе, т.е. к шуму “примешан” достаточно сильный тональный сигнал, и обе звуковые посылки (<шум> и <сигнал+шум>) без труда отличимы друг от друга. Во второй (тренировочной) части задача испытуемого состоит в подборе пороговой интенсивности тональной добавки и достижении асимптотического уровня обнаружения тонального сигнала. Стратегия работы испытуемого в тренировочной серии опыта и ее задачи подробно описаны в учебном задании, посвященном методу “Да-Нет”.

Для оптимизации тренировочного процесса при прослушивании стимулов испытуемый может включить режим “Подсказки”, когда перед каждой пробой указывается, какой из стимулов был сигнальным.

По окончании пробы в течение 3—4-секундного межпробного интервала (испытуемый сам подбирает его величину в тренировочной серии) испытуемый должен решить, какой стимул в паре (первый или второй) был сигнальным и дать ответ, нажимая на клавиши <1> или <2> цифровой клавиатуры, соответственно.

Опыт включает 400 проб: в 200 пробах на первом месте в паре предъявляется сигнальный стимул, в других 200 — пустой. Место сигнального стимула в паре меняется в квази-случайном порядке. После 200 проб делается перерыв.

После опыта целесообразно записать хотя бы краткий самоотчет, в котором стоит отметить свои наблюдения над особенностями стимуляции, своими переживаниями по ходу опыта, применявшимися способами выбора ответа и их изменениями в ходе опыта, если они имели место.

2. Метод оценки. Процедура опыта.

Структура опыта в целом почти ничем не отличается от изложенной выше для метода 2АВВ. В инструкции испытуемому подчеркивается, что после окончания каждой пробы в период межстимульного интервала необходимо оценить степень своей уверенности в наличии сигнала в данной пробе, используя 5-балльную шкалу оценок: *<5> — “точно, был сигнал, 100% уверенности”; <4> — “скорее всего, это был сигнал, 75% уверенности”; <3> — “то ли сигнал, то ли шум, 50% уверенности”; <2> — “скорее всего, это был шум, 25% уверенности”; <1> — “уверен в том, что это был шум, 0% уверенности”.* Ответ дается нажатием соответствующих клавиш на цифровой клавиатуре. Очень важно, чтобы в ходе ознакомительной серии испытуемый хорошо понял инструкцию и научился быстро и точно нажимать на нужные клавиши.

Опыт включает 500 проб: 250 сигнальных стимулов и 250 пустых или шумовых. Место сигнального стимула в последовательности проб меняется в квази-случайном порядке. В середине опыта делается перерыв.

После окончания опыта стоит также записать самоотчет.

Обработка результатов

Обработка результатов опыта 2АВВ состоит в следующем:

1.После окончания эксперимента студент получает распечатку результатов, где представлены вероятности всех 4-х типов исходов: p(H), p(FA), p(CR), p(O) и p(C). Результаты можно и переписать непосредственно из файла данных — это обычный ASCII-файл, имя которого соответствует фамилии студента по-латыни, а расширение — *abb,* например *sokolova.abb.*

Уточним, что при обработке данных компьютерная программа считала правильный ответ на стимул 1 — попаданием, правильный ответ на стимул 2 — правильным отрицательным ответом, ошибку на стимул 1 — пропуском, а ошибку на стимул 2 — ложной тревогой.

2. Далее необходимо провести проверку результатов опыта на *несмещенность*, т.е. на равенство p(H) и p(CR). Это делается следующим образом с помощью статистического критерия χ_2 (хи-квадрат):

а) вычисляются “ожидаемые” значения вероятностей правильных и ложных ответов:

$$P^* = \frac{p(H) + p(CR)}{2} \;;\; Q^* = 1 - P^* ;$$

б) вычисляется полученное в эксперименте значение $\chi^2_{\text{эксп.}}$;

$$\chi^2_{'\text{эксп.}} = N_S \frac{(p(H) - P^*)^2}{2} + N_N \frac{(p(CR) - P^*)^2}{2} +$$

$$+ N_S \frac{(p(O) - Q^*)^2}{Q^*} + N_N \frac{(p(FA) - Q^*)^2}{Q^*} \quad ;$$

в) сравнивается полученное значение $\chi^2_{\text{эксп.}}$ с критической величиной χ^2 для двух степеней свободы при уровне значимости a = 0,05. В случае $\chi^2_{\text{эксп.}} < \chi^2$ результаты эксперимента признаются несмещенными.

5. Вычисляется среднее значение P(C) по всей группе студентов (для расчета среднего значения следует взять данные не менее 10 человек).

6. Подсчитывается по индивидуальным, а затем и групповым данным *d'*.

Обработка результатов опыта МО проводится следующим образом:

1. В компьютерной распечатке приводятся условные вероятности отнесения сигнального и пустого стимулов к каждой из оценочных категорий, т.е. p(1)...p(5) и q(1)...q(5) и сводятся в табл. 1 ваших результатов, построенную аналогично табл. 5.

3. Последовательно суммируя p и q, вычисляются значения p(H) и p(FA) для каждого значения критерия (аналогично таблице 6) и вносятся в табл. 2 ваших результатов.

4. По данным табл. 2 на координатной бумаге строится кривая РХ.

Масштаб для построения РХ берется достаточно большим (не менее 100 мм на изменение вероятности от 0 до 1). Точки РХ соединяются на глазок плавной кривой.

5. Подсчитывается площадь под кривой РХ как мера сенсорной чувствительности или обнаружимости тонального сигнала на фоне шума.

6. Вычисляется средняя по всей группе площадь под кривой РХ.

7. Аналогично тому, как сравнивались вероятности p(H) и p(CR) при оценке несмещенности результатов

опыта 2ABB, определяется совпадение величин оценок сенсорной чувствительности в сериях 2ABB и ОУ у каждого испытуемого и по группе в целом. Для этого необходимо:

а) представить площадь под кривой РХ как теоретическую вероятность;

б) подсчитать V = 1 - U;

в) вычислить полученное в эксперименте значение $\chi^2_{\text{эксп.}}$:

$$\chi^2 = N\frac{[P(C)-U]^2}{U} + N\frac{[P(NC)-V]^2}{V},$$

где N — число измерений в серии 2ABB; P(C) и P(NC) — оценка вероятности правильных и неправильных ответов, соответственно;

г) сравнить полученное значение $\chi^2_{\text{эксп.}}$ с критическим значением $\chi^2_{\text{эксп.}}$ при 1 степени свободы и уровне значимости $\alpha = 0{,}01$.

8. По результатам серии ОУ построить кривую РХ в двойных нормальных координатах и вычислить d' по каждой точке РХ.

9. Сопоставить значения d', полученные в опыте 2ABB, с каждым из значений d' в опыте ОУ.

Обсуждение результатов

1) Сопоставить и вынести суждения о достоинствах и недостатках каждого из использовавшихся в задании методов при решении задачи оценки сенсорной чувствительности.

2) Если в серии 2ABB был получен смещенный случай, попытаться дать ему возможные объяснения, проанализировав тактику работы испытуемого (на основе самоотчета).

3) Сравнить полученное соотношение d'_{2ABB} и $d'_{ОУ}$ с теоретически ожидаемым. В случае, если указанное соотношение окажется не постоянным, попытаться дать объяснение этому факту, проанализировав соответствие результатов эксперимента исходным допущениям.

Литература
Основная

1. *Бардин К.В.* Проблема порогов чувствительности и психофизические методы. М.: Наука, 1976.

2. Проблемы и методы психофизики / Под ред. А.Г.Асмолова, М.Б.Михалевской М.: Изд-во Моск. ун-та, 1974. С. 145—169.

3. Хрестоматия по ощущению и восприятию / Под ред. Ю.Б.Гиппенрейтер, М.Б.Михалевской. М.: Изд-во Моск. ун-та, 1975. С. 233—248.

Дополнительная

1. *Иган Дж.* Теория обнаружения сигнала и анализ рабочих характеристик. М.: Наука, 1983. С. 17—83.

2. *Green. D.M., Swets, J.A.* Signal detection theory and psychophysics. N.Y.: Wiley, 1966.

3. *Gescheider. G.A.* Psychophysics: Method, theory and application. Hillsdale, NJ: Erlbaum, 1985.

Приложение 1

Дополнительные сведения о критериях принятия решения[1]

Критерий принятия решения в психофизической теории обнаружения сигнала (ТОС) - характеристика одной из двух основных составляющих процесса обнаружения сигнала: процесса принятия решения о характере стимульного воздействия. Это понятие используется как для описания оптимальности решения наблюдателя в принципе (так называемый критерий оптимальности), так и для оценки реально используемой им стратегии решения сенсорной задачи (так называемый критерий наблюдателя).

<u>Критерий оптимальности</u> (качества) решения (КрО): в современной психофизике - мера (показатель) эффек-

---

[1] Приведенный ниже материал представляет собой неопубликованную статью И.Г. Скотниковой для “Психофизической энциклопедии”.

тивности решения, которая детерминируется его целью, и в соответствии с ней отражает также предпочтительность способов и результатов деятельности. КрО в явном виде задается инструкцией к задаче, сообщающей априорные вероятности предъявления сигнала и шума, а также стоимости каждого из 4-х исходов решения (попаданий, ложных тревог, правильных отрицаний и пропусков сигнала). Стоимости выражаются суммами выигрыша за верные решения и проигрыша за ошибочные. Вероятности и стоимости могут быть либо явно заданы испытуемому в инструкции, либо усвоены им самостоятельно в ходе решения здачи, основываясь на субъективной оценке вероятности предъявления сигнала и информации от экспериментатора о качестве его работы. Таким образом, *соотношение априорных вероятностей и стоимостей* определяет отличие видов КрО друг от друга.

Существует следующая классификация КрО:

1. Критерий, введенный в рамках ТОС: цель решения — максимизировать выигрыш. Стоимости 4-х исходов решения могут быть любыми, причем для верных ответов они положительные (либо нулевые), для ошибочных — отрицательные (либо нулевые). Наблюдатель должен учитывать все их для вынесения оптимального (максимально выигрышного) суждения. Численное значение КрО может быть выражено следующей формулой:

$$K = P(S)P(Y/S)C(Y/C) - P(S)P(N/S)C(N/S) + P(N)P(N/N)C(N/N) - P(N)P(Y/N)C(Y/N),$$

где K — критерий оптимальности; P(S) и P(N) — априорные вероятности появления сигнала и шума, соответственно; P(Y/S) и P(Y/N) — вероятности попаданий и ложных тревог соответственно; P(N/S) и P(N/N) — вероятности пропусков и правильных отрицаний, соответственно; C(Y/S) и C(Y/N) — стоимости попаданий и ложных тревог, соответственно; C(N/S) и C(N/N) — стоимости пропусков и правильных отрицаний, соответственно.

2. Критерий Байеса: цель решения - минимизировать проигрыш (средний риск). Стоимости верных ответов − нуле-

вые, ошибочных – отрицательные, что побуждает наблюдателя ориентироваться преимущественно на ошибки. Критерии ТОС и Байеса близки по смыслу (они — игровые) и по цели решения (максимизировать выигрыш и минимизировать проигрыш) и поэтому обычно объединяются в общий класс КрО. Как частные случаи КрО по ТОС и Байесу рассматриваются:

Критерий Котельникова (идеального наблюдателя): цель решения — минимизировать суммарную ошибку наблюдения. Стоимости верных ответов — нулевые, ошибочных — равные отрицательные (т.е. разновидность критерия Байеса). Наблюдатель оценивает суммарную ошибку пропусков и ложных тревог и поддерживает ее на постоянном минимальном уровне.

Критерий Неймана-Пирсона: цель решения — минимизировать вероятность ошибок одного рода (обычно — пропусков сигнала) при фиксации вероятности ошибок другого рода (обычно — ложных тревог). В таком типичном случае стоимость ложных тревог максимальна , в сравнении со стоимостями остальных 3-х исходов (которые либо нулевые, либо равные). Информацию о стоимостях и априорных вероятностях сигнала и шума наблюдатель может достоверно не знать (лишь допускать ее). При этом использование данного критерия оптимально. Так обычно работает необученный наблюдатель, поддерживая уровень ложных тревог постоянным, независимо от характера сигнального распределения.

Минимаксный критерий: цель решения – минимизировать максимальные ошибки обоего рода. Оптимален в случаях, когда наблюдатель не знает точно априорные вероятности сигнала и шума и ориентируется на свой опыт, усвоенный в ходе экспериментов, стремясь уравнять вероятности обеих ошибок.

3. Дж.Иган кроме уже отмеченных выше критериев выделяет также критерий Зигерта: цель решения — максимизировать процент правильных ответов. Стоимости верных ответов равны стоимостям ошибочных, поэтому наблюдатель максимизирует как ожидаемый выигрыш, так и долю правильных ответов.

4. Ю.М.Забродин выделяет два рода КрО. Критерий 1-го рода: цель решения - минимизировать субъективную неопределенность относительно входной информации. Это информационный критерий и наиболее естественный для человека: вынесение суждения на основе априорной информации (физических и вероятностных свойств сигналов) и апостериорной (вероятностей ответов). Критерий 2-го рода: цель решения – достичь максимально устойчивой деятельности. Это игровой критерий: ориентация на стоимости ответов. Обычно необученный наблюдатель вначале использует критерий первого рода, далее, усвоив информацию о стоимостях, — критерий второго рода. Субъективное представление о стоимостях определяет выбор наблюдателем вида КрО в рамках критерия 1-го или 2-го рода.

Критерий наблюдателя (КрН) – критическое значение сенсорного впечатления в ряду наблюдений (разделяющая граница на сенсорной оси, используемая субъектом для сравнения с каждым наблюдением с целью выбора ответа по результату этого сравнения в отличие от критерия оптимальности, устанавливаемого на оси отношения правдоподобия). Если сенсорный эффект данного наблюдения меньше критического, выносится ответ “Нет” (нет сигнала), если больше, то ответ “Да” (есть сигнал).

Ряд теоретических моделей по-разному описывают правила выбора субъектом КрН и способы расчета его числовых значений на основании вероятности ответов.

В рамках ТОС правило принятия решения о наличии сигнала основывается на оценке *отношения правдоподобия*, т.е. отношения плотности вероятности того, что данное сенсорное событие (Xs) вызвано предъявлением шума (N) к плотности вероятности того, что оно вызвано предъявлением сигнала (S): (f(X/N) / f(X/S). Отношение правдоподобия сравнивается с критическим его значением C_0, которое выполняет функцию КрН. Субъект выбирает значение КрН (“Бетта”) на основании априорной информации о вероятностях сигнала и шума и о стоимостях ответов. Теоретическое (задаваемое этими характеристиками) значение КрН имеет вид:

β= P(N)/P(S) C(N/N)-C(Y/N)/C(Y/S)-C(N/S).

Эмпирическое значение КрН графически определяется как тангенс угла наклона касательной к данной точке РХ (что соответствует производной в этой точке) и аналитически рассчитывается на основании вероятностей попаданий и ложных тревог, для которых находятся табличные значения соответствующих им плотностей вероятности нормального распределения. Существуют сводные таблицы значений β для любой пары вероятностей попадания и ложных тревог. В соответствии с классической ТОС предполагается, что положение критерия тренированного наблюдателя полагается неизменным в ходе опыта. В ряде других моделей описывается различного рода динамика положения критерия.

Приложение 2

Краткое описание программы yes_no.exe

Настоящая программа предназначена для проведения психофизических экспериментов по обнаружению зрительных сигналов с использованием метода “Да-Нет”. Стимульный паттерн представляет собой две колонки текстовых элементов (буквы, цифры или знаки псевдографики) по 3 элемента в колонке. Ниже на рис. 9 показан один из вариантов такого паттерна.

```
R   R
R   R  — вариант  сигнального стимула
L   R

R   R
R   R  — вариант несигнального стимула
R   R
```

Рис. 9. Сигнальный и несигнальный стимулы

Как явствует из приведенного рисунка, *сигнальный и несигнальный* стимулы отличаются только одним символом (в данном случае - это L внизу левого верхнего столбца). В программе имеется некоторый стандартный набор используемых символов, кроме того, экспериментатор может самостоятельно дополнять этот набор по собственному желанию, дописывая нужные символы в текстовый файл под названием stimul.dat. Местоположение этого сигнального элемента меняется от пробы к пробе — он может появляться *в случайном порядке* на любом из 6 знакомест. В соответствии с процедурой метода "Да-Нет" сигнальные и несигнальные пробы предъявляются в случайном порядке, причем вероятность предъявления сигнальной пробы в ряде проб, составляющих одну экспериментальную серию, может задаваться экспериментатором.

При подготовке опыта экспериментатор должен установить *в режиме меню* необходимые параметры стимуляции и ответа испытуемого.

Программа снабжена подробным описанием каждого пункта меню, помощь вызывается нажатием функциональной клавиши F1. После задания всех параметров они запоминаются программой и при запуске очередной серии их изменение не обязательно. Передвижение по пунктам меню осуществляется клавишами управления курсора (вверх - вниз). По ряду технических причин целесообразно изменять параметры, двигаясь по пунктам меню *сверху вниз.* При раскрытии ряда пунктов меню выбор необходимого параметра осуществляется этими же клавишами. Изменение параметров других пунктов меню осуществляется введением с цифровой клавиатуры необходимого числа. Подчеркнем, что введение фамилии и имени является обязательным, поскольку они задают имя файла результатов. В том случае, если экспериментатор забыл их задать, то при попытке начать опыт, программа возвращает его к этим пунктам. Быстрый переход к пункту "Начать эксперимент" осуществляется нажатием клавиши Tab. Как правило, в ходе опыта (особенно в тренировочной его части) экспериментатор наиболее часто из-меняет три основных параметра стимуляции – *длительность предъявления* стимула, *межстимульный интервал (МСИ)* и *количество* стимулов. Первые два параметра задаются в миллисекундах (1000 мс = 1 с). О выборе количества стимулов говорилось выше, а о временных параметрах следует сказать особо. Величина МСИ выбирается из соображений удобства и комфортности для испытуемого, но ее не следует делать слишком короткой или слишком длинной. Как правило, она выбирается в диапазоне от 1.5 до 2.5 с. Особо остановимся на описании следующих параметров:

– “Длительность предъявления маскера.” После окончания предъявления стимула в программе предусмотрена (хотя и необязательно) возможность предъявления “маскера” специального стимула, “затирающего”, маскирующего сетчаточный образ основного стимула, и тем самым затрудняющего его восприятие. Использование маскера может быть полезным в случае высокой чувствительности испытуемого для усложнения задачи обнаружения сигнала. Как и в случае выбора стимульных символов, маскером может быть любой текстовый символ. Четыре комбинации стимул - маскер представлены в меню “Вид стимуляции”. Если Вы не используете маскер, то установите длительность его предъявления равной нулю.

– “Расстояние от стимула до линии (позиции)”. Этот параметр определяет, на сколько пустых символьных интервалов (позиций) буквы отстоят от вертикальной линии в центре экрана. Понятно, что чем больше расстояние между столбцами, тем труднее обнаруживать сигнальную букву. Стандартным (заданным в программе по умолчанию) значением является 9 позиций.

– “Для ответа использовать клавиши...” Этот пункт меню позволяет выбрать два варианта ответной реакции испытуемого: использовать клавиши Shift (левый и правый) или клавиши управления курсором (налево — направо).

– “Отвечать “ДА” клавишей...” В этом пункте меню можно выбрать ту клавишу, которая будет соответствовать ответу “да”.

– <u>С какого члена последовательности начать опыт? Количество предъявлений.</u> Эти пункты меню устанавливают количество проб в каждой серии и начальный член случайной последовательности предъявления сигнальных и шумовых проб. Последнее позволяет исключить запоминание испытуемым нескольких первых чередований сигнальных и шумовых стимулов в случайной последовательности проб.

После окончания каждой серии эксперимента рассчитываются все необходимые показатели обнаружения сигнала. Полученные результаты выводятся на экран компьютера, где по порядку (от первой серии до последней) они представлены в виде таблицы, являющейся протоколом вашего эксперимента. Ниже приведен фрагмент такого протокола и даны пояснения используемым показателям. В случае необходимости вы всегда сможете обратиться к соответствующему файлу на диске, в котором хранится ваш протокол.

Используемые обозначения:

t(st) — длительность стимула в миллисекундах;

P(S) — априорная вероятность предъявления сигнальной пробы;

P(H) — вероятность правильных обнаружений (попаданий);

P(FA) — вероятность ложных тревог;

d' — индекс сенсорной чувствительности;

β — индекс критерия принятия решения (отношения правдоподобия);

ВР — среднее время реакции в данной серии, в миллисекундах;

σ ВР — стандартное отклонение ВР, в миллисекундах;

Hit, CR, M, FA - количество попаданий, правильных отказов, пропусков и ложных тревог, соответственно.

Таблица 8

Фрагмент протокола эксперимента, выводимого программой на экран монитора после окончания серии

№ сер.	t(st)	P(S)	P(H)	P(FA)	d'	β	BP	σ BP	Hit	CR	Miss	FA
1	300	0.5	0.89	0.1	3.01	11.01	984	141	89	90	11	10
2	200	0.5	0.81	0.15	2.96	1.28	726	125	81	85	19	15
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
11	150	0.5	0.75	0.02	2.01	1.00	798	123	75	98	25	2

ЧАСТЬ II
ОДНОМЕРНОЕ ШКАЛИРОВАНИЕ

Глава 1. МЕТОД БАЛЛЬНЫХ ОЦЕНОК

Излагаемые в данной главе процедуры (ranking procedures, Шарф, 1975) относятся к наиболее распространенным методам *порядкового* шкалирования. В отечественной литературе они получили общее название *метода балльных оценок*, хотя, как будет видно из излагаемого материала, они не ограничиваются только числовыми оценками стимулов. В некоторых случаях метод балльных оценок может дать более “сильную” шкалу, чем порядковая. Однако это счастливое исключение из правила, связанное более с измерительным опытом наблюдателя и характеристиками оцениваемых объектов, чем с особенностями самой измерительной процедуры.

В первой части этой главы рассматриваются основные принципы метода балльной оценки и приводятся наиболее распространенные алгоритмы построения шкал балльных оценок. Во второй части анализируются основные артефакты измерения, связанные с построением шкалы балльных оценок. Будут рассмотрены также некоторые специальные условия, которые рекомендуется соблюдать при использовании метода балльных оценок.

Из всех методов психологических измерений, в которых используются оценочные суждения человека, процедура шкалирования, основанная на балльных оценках, наиболее популярна в силу своей простоты. Распространенность этого метода связана с прикладными разделами психологии, но не менее широко он используется и в академических исследованиях, например, в психодиагностике при оценке различий испытуемых или в психофизике при психологической оценке стимулов.

Наиболее распространенные разновидности метода балльных оценок делятся на пять больших классов: *класс числовых методов, графических и шкалирование с исполь-*

зованием стандартов, кумулятивных и методов вынужденного выбора (Гилфорд, 1954). Все эти классы связаны с распределением объектов (стимулов) либо вдоль непрерывного континуума, либо в виде упорядоченных дискретных категорий. Все методы похожи тем, что конечным результатом является приписывание чисел стимулам в соответствии с порядком их распределения по континууму, а различаются они либо процедурой распределения стимула, либо способом различения стимулов и количеством вспомогательных операций, необходимых испытуемому. Существуют и некоторые другие аспекты, по которым они различаются, но в связи с их частным характером эти аспекты будут рассмотрены по ходу описания каждого класса в отдельности. В данной главе будут описаны первые три класса методов, поскольку они наиболее часто используются в психологических измерениях.

§ 1. Графические шкалы

Наиболее распространенным типом шкалы балльных оценок является, вероятно, графическая шкала. В общем случае она представляет собой прямую линию, на которой определенным образом размечены признаки, характеризующие исследуемый класс объектов-стимулов. Линия может быть разделена на отрезки или непрерывной. Если она разделена на отрезки, то число отрезков может быть различным. Она может быть расположена горизонтально или вертикально. Пример непрерывной графической шкалы для балльной оценки скорости мышления отдельных индивидов приведены на рис. 1.

Крайне медленно мыслит	*Инертный тугодум*	*Мыслит с обычной скоростью*	*Живой ум*	*Чрезвычайно быстрое мышление*

Рис. 1. Пример графической шкалы для оценки скорости мышления

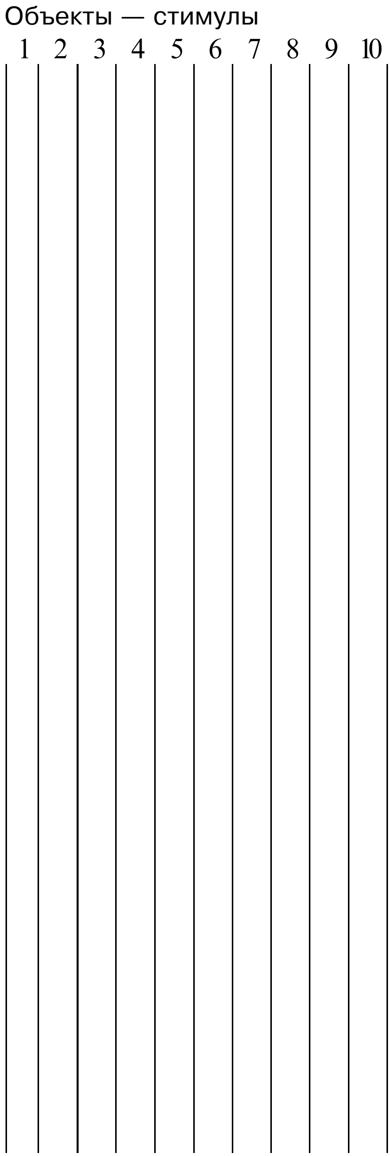
Испытуемый в этом случае выносит суждение в форме отметки на графической шкале. Признаки, расположенные вдоль шкалы, помогают ему сделать суждение более точным.

Параллельные графические шкалы. Другая форма графической шкалы балльных оценок, названная шкалой балльных оценок поведения, была разработана Чемпнеем (1940) для оценки некоторых характеристик окружающей ребенка домашней среды. Пример такой шкалы приведен на рис. 2.

В данном примере инструкция испытуемому была такова: “Оцените родительское стремление проявить сверхзаботу о детском благополучии. Действительно ли родители паникуют в зависимости от степени важности ситуации, или есть родители относительно спокойные, холодные или беззаботные к своему ребенку даже в критических ситуациях?”. Кроме того, подчеркивалось, что поведение родителей рассматривается независимо от стоящих за ним мотивов, и в оценку включается только то поведение, которое потенциально направлено на ребенка и которое касается его физического и психического здоровья и комфорта.

Основная особенность этой шкалы состоит в том, что линии балльных оценок располагаются вдоль *вертикальной линии*. Дело в том, что на горизонтальной линии можно предусмотреть место только для очень коротенького описания признака. Подробное многословное описание здесь уже не поместится. Кроме того, на горизонтальной линии признак труднее локализовать в определенной точке, он оказывается как бы распространен вдоль линии, и поэтому точное положение его на шкале не совсем ясно. При использовании вертикальных линий эти трудности легко устранить. Признаки могут быть достаточно подробными для того, чтобы быть более значащими, в то же время их можно точнее локализовать в точках шкалы.

Еще одна положительная особенность этой шкалы заключается в том, что на каждой странице оценивается *только одна характеристика* поведения (в данном примере — “озабоченность — беспечность”).



Предрасположен к строгости, иррациональной тревожности, в значительной степени на иррациональной основе.

Постоянное тревожное напряжение по поводу ребенка, но скорее “нервное”, чем паника.

Склонен видеть опасность там, где ее реально нет.

Проявляет значительное беспокойство, но редко теряет разумный контроль над собой.

Заботливый, но склонен преуменьшать опасность. Очень часто сверхвнимателен, но ситуацию оценивает правильно, не теряет перспективы.

Редко обеспокоен или озабочен теми аспектами поведения, которые выходят за пределы непосредственной ситуации и ответственности. Социальная установка похожа на установку учителя или няни.

Беззаботный и, по-видимому, беспечный даже при важных делах. Настолько беззаботный, что оказывается невнимательным и безответственным.

Рис. 2. Пример графической шкалы балльных оценок поведения родителей по характеристике “озабоченный — беспечный” (Гилфорд, 1954)

Общеизвестно, что более адекватные оценки дают процедуры, в которых испытуемый имеет возможность оценить всех членов группы *по одной* характеристике, а потом уже переходить к другой. Однако эта процедура дает удовлетворительные результаты, если в ней контролируется хорошо известный “гало-эффект” (см. ниже). Отметим также, что с помощью параллельных линий балльных оценок одной и той же характеристики можно сделать оценки сразу для *многих* объектов.

<u>Общие рекомендации к построению графических шкал.</u> Есть определенные эмпирические правила, соблюдение которых способствует эффективности графических балльных оценок (Гилфорд, 1954). Не все из них достаточно бесспорны и убедительны, но исследователю нужно о них помнить:

1. Все объекты должны быть оценены по одной характеристике, и только потом можно переходить к следующей характеристике.

2. Линии должны быть по крайней мере 15 см длиной, но не намного длиннее. Линия должна быть достаточно длинной, чтобы учитывать самые точные количественные различия, которые могут дать испытуемые. Но при очень длинных линиях единство континуума для испытуемого прерывается. Длинные линии часто заставляют испытуемого локально сгущать оценки, а не распределять их непрерывно.

3. Линии не должны иметь разрывов и делений. Но единого мнения о том, какую из двух видов линий использовать, непрерывную или дискретную, нет. Непрерывная линия подчеркивает непрерывность шкалируемой характеристики. Дискретная линия может предполагать разрывность или скачкообразные качественные изменения оцениваемой переменной. Непрерывная линия может быть разделена на любое число единиц, и деления могут быть размещены в соответствии с предпочтением испытуемого.

4. Для “неиспорченных” и необученных испытуемых “хорошая” оценка обычно связана с началом линии слева или

сверху. В вертикальных шкалах “хорошую” оценку располагают вверху — это естественно для всех. В горизонтальных же шкалах наличие “хорошей” оценки противоречит обычной практике математической системы координат. Но, тем не менее, испытуемые обычно предпочитают помещать положительные значения оцениваемой характеристики в начале линии, слева.

5. Описательные фразы и признаки должны быть сконцентрированы по возможности у точек на шкале. Это очень легко сделать для вертикальных шкал. Для горизонтальных шкал полезно использовать слова, располагающиеся в колонке одно над другим.

6. Необходимые признаки обычно равномерно расставляются вдоль линии; но это можно делать, только если они одинаково различны. В противном случае сами признаки должны быть прошкалированы какой-то отдельной психологической процедурой и тогда их локализация будет обусловливаться уже этой шкалой. Иногда между признаками промежутки специально искажаются, чтобы противодействовать общим смещениям (байесам) в балльных оценках. Например, чтобы противодействовать ошибке “смягчения” (см. ниже), признаки на предпочитаемой стороне шкалы располагают с более широкими интервалами, чем признаки на непредпочитаемой стороне. Чтобы противодействовать тенденции образовывать сгущения балльных оценок к середине шкалы (эффект центрации), промежутки между средними признаками можно немного увеличить.

7. Конечные признаки не должны быть такими крайними по содержанию, что испытуемые очевидно никогда не будут ими пользоваться. Положение конечных признаков должно быть близко к концам линии.

8. В случае биполярных характеристик нейтральный или индифферентный признак находится обычно в центре линии, если не вводятся модификации, например, типа правила 6.

9. В процессе шкалирования можно использовать трафарет, который разделяет каждую линию на секции, где, в свою очередь, могут использоваться числовые оценки. Деления могут быть неравными, они могут быть изменены с

тем, чтобы помочь противодействовать систематическим байесам в балльных оценках или нормализовать распределения шкал.

Оценка графических шкал. У графических шкал много достоинств и сравнительно мало недостатков. Среди наиболее существенных преимуществ — простота и легкая управляемость. Эти шкалы интересны и не требуют сильной дополнительной мотивации, процедура шкалирования быстро выполняется испытуемым, не требует от него числовых операций. С точки зрения теории измерения, графическая шкала обеспечивает возможность такого точного различения, на которое испытуемый вообще способен, т.е. графическая шкала может обладать “силой” шкалы интервалов или отношений, хотя чаще всего она представляет собой шкалу порядка.

§ 2. Числовое шкалирование

В числовом методе построения шкалы балльных оценок испытуемому дается последовательность определенных чисел (баллов или рангов) и он приписывает каждому стимулу соответствующее число из ряда. Пример такой шкалы, которую использовал Гилфорд (1954) для получения балльных оценок аффективных характеристик цветов и запахов, приводится ниже:

10— Невообразимо приятный
9— Наиболее приятный
8— Очень приятный
7— Умеренно приятный
6— Чуть-чуть приятный
5— Безразличный
4— Чуть неприятный
3— Умеренно неприятный
2— Очень неприятный
1— Крайне неприятный
0— Невообразимо неприятный

Некоторые “числовые” шкалы, например, шкала успеваемости, на самом деле основываются на описательных суждениях типа:

Отлично

Хорошо

Удовлетворительно

Плохо

Очень плохо

Затем этим прилагательным экспериментатор приписывает числа, например, от 5 до 1. При такой процедуре предполагается, что *психологические интервалы* между прилагательными равны, но с точки зрения их уточнения лучше, чтобы сам испытуемый непосредственно пользовался этими числами (Торгерсон, 1958).

<u>Некоторые проблемы числовых шкал:</u>

1. Использование отрицательных чисел. Шкала аффектов и шкала успеваемости, рассмотренные выше, являются биполярными. Континуум представляет собой изменения в двух противоположных направлениях. По этой причине некоторые исследователи помещают ноль в нейтральной или средней категории, а отрицательные числа — ниже его. Это более естественно для того, кто знаком с алгеброй, но может быть неестественным для менее образованных испытуемых. Другая опасность состоит в том, что биполярность может создать впечатление о разрыве в нулевой точке шкалы и тем самым нарушить предполагаемую непрерывность. Наличие отрицательных чисел, таким образом, может иметь сложности для исследователя. По этим причинам использование отрицательных балльных оценок не рекомендуется (Гилфорд, 1954).

2. “Заякоривание” аффективной шкалы. Может показаться, что два крайних прилагательных в первом примере бесполезны и что вряд ли кто-нибудь из испытуемых будет пользоваться крайними категориями. Вообще, избегание крайних категорий, которые испытуемый заведомо не использует, можно считать хорошей тактикой. Однако имеются два аргумента в пользу того, чтобы включать именно такие конечные прилагательные. Один состоит в том, что некоторые испытуемые на самом деле все-таки используют даже самые крайние категории. Кроме того, испытуемый всегда может столкнуться с таким стимулом, который явно соответствует более крайней категории, чем любой из тех, что заносились в категорию 9 или 1. Если бы не было более крайних категорий, испытуемый вынужденно оценивал бы этот стимул как равный другим, хотя он явно видит их неравенство. Таким образом, конечные категории могут служить для выхода из крайних положений, которые иногда возникают. Другой аргумент состоит в том, что крайние категории являются "якорями" для всей шкалы. Показано, что добавление такой категории к одному из двух концов помогает расширить (т.е. увеличить дисперсию) исходное распределение балльных оценок в направлении этой категории (Хант и Фолькмен, 1977). В любом случае у испытуемых имеется общая тенденция *избегать конечных категорий* (и одновременно с этим сдвигать все оценки немного по направлению к середине ряда). Если категории 0 и 10 не были включены, испытуемые будут иметь тенденцию избегать категории 1 и 9, и, таким образом укорачивать ряд балльных оценок. Итак, если исследователь хочет иметь эффективную шкалу из девяти точек, он должен обеспечить возможность расширить выход за эти 9 точек, а иначе он может в конце получить шкалу меньшую, чем из 9 точек.

Оценка числовых шкал. Для испытуемого числовые шкалы — самые легкие по вынесению суждений, а для экспериментатора — самые простые с точки зрения обработки результатов. Если испытуемый работает добросовестно, если свойства чисел можно в принципе применять к наблюдаемым феноменам, то балльные оценки сами по себе оказываются соответствующими "сильной" шкале. Эмпирическая проверка числовых балльных оценок на свойства шкалы интервалов и шкалы отношений сделана в ряде работ (Соколов и др., 1978; Ратанова, 1972). Строгие методы проверки этих свойств для данных, полученных числовым методом балльных оценок, можно найти у Гилфорда (1954).

§ 3. Шкалирование по стандартной шкале

Особенность этого типа шкал состоит в том, что испытуемому предоставляется некоторый *набор стандартов* того же вида, что и оцениваемые стимулы. Лучшим примером такой шкалы служат шкалы для оценивания свойств почерка. Эти шкалы снабжены отдельными образцами, которые заранее проградуированы по “сильной” шкале каллиграфического качества — например, таким методом, как равновоспринимаемые интервалы или методом парных сравнений. При наличии шкалы стандартов новый образец почерка может быть легко уравнен с одним из стандартов или оценен как находящийся между двумя стандартами.

Другой формой этой шкалы является использование в качестве стандартов *стандартных оценок* вместо отградуированных образцов. Примером такой процедуры служит методика подбора пары к образцу, которая была разработана Хартшерном и Мэем (1929) в связи с изучением характера.

Метод подбора пары к образцу. Построение набора (у Хартшерна и Мэя — вербальных портретов) по выбранной характеристике состоит из нескольких этапов. Во-первых, было собрано большое количество утверждений, имеющих отношение к проявлениям данной черты характера. Каждое утверждение было записано на отдельной карточке, а карточки были проранжированы группой экспертов. Были составлены 10 описаний или портретов. Каждое состояло из утверждений, имеющих приблизительно один и тот же средний ранг. Портреты снова были проранжированы 48 наблюдателями, и так были получены для них стандартные шкальные оценки. Например, портрет со шкальной оценкой “7” по такой черте характера, как “полезность людям”, имеет следующую форму:

“X — всегда заботится о людях, старается быть полезным окружающим, не ожидая, когда его попросят об этом. При случае он готов помочь кому-либо в опасности; свои собственные интересы и гордость у него на втором плане; он мало озабочен отдаленными нуждами, особенно, если они не слишком значительны или серьезны".

При использовании портретов испытуемый читает определенное описание и затем называет всех индивидов, которых, как ему кажется, это касается. Один и тот же индивид может быть назван в связи с более, чем одним, портретом. Окончательная балльная оценка есть медиана всех оценок портретов, которые давались всеми испытуемыми.

Оценка процедур с использованием шкалы стандартов. Основное преимущество этих методов в том, что создаются более или менее постоянные эталоны, которые служат уже объективными вехами, помогая испытуемому стабилизировать оценки. Если есть хороший набор объективных стандартов, который широко применяется (как в случае шкал почерков), то метод шкалирования со стандартами имеет большое преимущество в стабильности результатов.

§ 4. Проблемы, связанные с построением шкал балльных оценок

Постоянные ошибки и их контроль. Использование балльных оценок основывается на предположении, что человек-наблюдатель является хорошим инструментом количественного наблюдения, что он способен делать точные и объективные суждения. Тем не менее, хотя мы и предполагаем возможность вынесения количественных суждений, мы должны всегда быть бдительными к влиянию предпочтений (байесов) в этих суждениях. Следствием этого влияния могут быть систематические ошибки в суждениях испытуемых. Рассмотрим некоторые наиболее распространенные байесы в процедурах метода балльных оценок.

Ошибки "смягчения" суждений. Многие испытуемые имеют тенденцию оценивать то, что они хорошо знают или то, что чаще встречается, *выше, чем следует*. Это — систематическая ошибка, которая не зависит от шкалируемого признака. Некоторые самокритичные испытуемые, которые отдают себе отчет в этой слабости, могут, в результате, удариться в другую крайность и давать оценки *ниже, чем следует.* Для описания таких отклонений и используется термин “ошибка смягчения”, применимый к общей постоянной тенденции испытуемого оценивать объекты шкалирования слишком высоко или слишком низко. При занижении оценки постоянная ошибка называется ошибкой негативного смягчения. Так как положительная ошибка смягчения является намного более общей, некоторые исследователи пытались предвосхитить ее и изменить шкалу так, чтобы нейтрализовать ошибку. Примером такой модификации шкалы успеваемости, рассмотренной выше, может служить следующая шкала:

Плохая	Средняя	Хорошая	Очень хорошая	Блестящая

Рис. 2. Пример модифицированной шкалы успеваемости

На этой шкале большинство признаков имеет благоприятное значение.

Можно предвидеть, что средняя балльная оценка расположится где-то рядом с признаком “хорошее” и распределение будет симметрично относительно этой точки.

Ошибка центрации. Одной из причин ошибки центрации, или, как ее называют, центральной тенденции, является та, что испытуемый реже дает крайние утверждения и, таким образом, смещает оцениваемые объекты-стимулы в направлении к середине всей группы. Это особенно характерно для балльных оценок таких объектов, о которых эксперты-испытуемые *знают не очень много.* По этой причине в связи с графическими шкалами давалась рекомендация располагать описательные фразы в середине шкалы с бульшими промежутками, чем на краях. Подобным же образом в числовой шкале интенсивность описательных прилагательных может быть установлена так, чтобы значения у концов шкалы больше различались между собой, чем значения у центра при одном и том же расстоянии между ними на линии. Это окажет противодействие ошибке центрации.

Влияние контекста. Ошибка центрации является частным случаем более общего типа ошибок, связанных с влиянием контекста на суждения испытуемого. Согласно традиционному подходу при шкалировании сенсорных и перцептивных объектов главный интерес исследователей сосредотачивался на получении оценок каждого стимула, величина которых, как предполагалось, определяется только наличным сенсорным впечатлением и не зависит от стимульного контекста. Первая брешь в этом подходе была пробита *Хелсоном*, разработавшим *теорию уровня адаптации* (УА). Согласно Хелсону (1975), оценку любой характеристики стимула (например, вес, яркость, размер) человек соотносит со своей субъективной шкалой, точнее говоря, с нейтральной точкой на этой субъективной шкале или точкой отсчета, названной Хелсоном уровнем адаптации. Значения стимулов, превышающие величину стимула, соответствующего уровню адаптации, оцениваются как более “тяжелые”, “яркие”, “большие”, а не достигающие этой величины — как более “легкие”, “тусклые”, “маленькие”. Хелсон считает, что уровень адаптации является суммарным результатом трех классов воздействия: 1) ряда стимулов, оцениваемых в данном эксперименте; 2) всех других стимулов, воздействовавших на человека во время измерения и составляющих контекст для первого класса стимулов и 3) стимулов, действовавших в прошлом на этого человека и оставивших след в его памяти. Уровень адаптации вычисляется как среднее геометрическое всех воздействовавших стимулов. В случае получения категориальной шкалы, когда испытуемый использует заданное число категорий для суждения о стимуле, грубая оценка уровня адаптации может быть получена путем вычисления среднего арифметического в средней или нейтральной категории (например, если используется 9 категорий в суждении о стимулах, то стимул, соответствующий среднему среди попавших в пятую категорию, характеризует уровень адаптации). Более точная оценка уровня адаптации получается на основе использования всех полученных в эксперименте данных и состоит в нахождении с помощью метода наименьших квадратов наилучшей аппроксимации полученной психофизической зависимости и определении по ней стимула, соответствующего нейтральной точке шкалы суждений.

Экспериментальное исследование влияния различных контекстов на простые перцептивные суждения проводилось *Пардуччи и Маршилл* (1961). Эти исследователи предъявляли испытуемым наборы линий с различным распределением их длин и просили их высказывать суждение о длине, приписывая каждой линии одну из шести категорий — от 1 (очень короткая) до 6 (очень длинная). Наборы линий различались по значению либо среднего арифметического длин линий, либо медианы, либо средней точки — средней между самой длинной и самой короткой линиями в ряду. Было обнаружено, что на центрирование шкалы суждений, отражавшихся в величине уровня адаптации, заметно влияет только изменение медианы в наборе линий. Изменение среднего арифметического ряда линий не оказывает существенного влияния на уровень адаптации.

В других работах Пардуччи (1974) было установленно, что форма психофизической функции, полученной при использовании категориальных суждений, зависит от характера распределения стимулов в используемом диапазоне стимуляции. Им было показано, что форма функции является более крутой для части стимульной области, где стимулы расположены более плотно в пространстве или предъявляются с большей частотой. Эти данные хорошо согласуются с гипотезой Пардуччи о том, что шкала суждений представляет собой компромисс между двумя различными тенденциями: 1) выносить суждение о длине линий, опираясь на перцептивное впечатление от стимула; 2) использовать различные категории суждений с равной частотой.

“Гало-эффект”. Постоянную ошибку, связанную с влиянием всей личности оцениваемого индивида на оценку отдельной черты характера (Уэлс, 1907), называют “гало-эффектом”. “Мы судим о наших близких с точки зрения общей умственной установки на них и эта умственная установка по отношению к личности, как к целому, преобладает над установкой в отношении отдельных черт характера", — говорил Торндайк (1920). Результатом "гало-эффекта" будет усиление балльной оценки любой характеристики, совпадающей с *общим впечатлением* от оцениваемых индивидов. Это делает балльные оценки некоторых характеристик менее валидными. Другим результатом этого байеса будет неверное количество положительных корреляций между оцениваемыми характеристиками. Из-за этого балльные оценки, в которых нет "гало-эффекта", некоторым образом сводятся на нет. "Гало-эффект" похож на ошибку стимула в психофизике. Он включает иррелевантный критерий, которым зашумляются суждения. Конечно, избежать полностью гало-ошибки невозможно, но опыт показывает, что вероятнее всего она обнаруживается в следующих случаях (Саймондз, 1925):

1. В признаках (характеристиках), которые трудно пояснить.
2. В экзотических, нетрадиционных признаках.
3. В недостаточно четко определенных характеристиках.
4. В признаках, включающих связи с другими людьми.
5. В характеристике, имеющей высокую моральную ценность. Это относится и к так называемым чертам характера.

Наилучший способ избежать "гало-эффекта" достигается при использовании метода графического шкалирования, где в каждом случае оценивается только одна характеристика.

<u>Логическая ошибка в балльной оценке.</u> Ошибка, вызванная тем, что эксперты или испытуемые дают одинаковые балльные оценки для тех характеристик, которые кажутся логически соотнесенными друг с другом, называется логической ошибкой (Ньюкомб, 1931). Так же, как и гало-эффект, эта ошибка искажает взаимосвязи характеристик, увеличивая эти связи, но по другой причине. В "гало-эффекте" это является следствием очевидной для испытуемого связанности отдельных личностных качеств, тогда как в логической ошибке это является следствием логической со-

гласованности различных характеристик, независимо от индивидов. Логической ошибки можно избежать, обращая внимание испытуемого на объективно наблюдаемые связи, а не на абстрактные логические или семантические совпадения характеристик.

Ошибки контраста. Под любой ошибкой контраста подразумевается тенденция испытуемого переоценивать других людей в противоположном направлении по сравнению с самим собой. Например, испытуемые, которые сами очень аккуратны, имеют тенденцию оценивать других как менее аккуратных, чем они есть на самом деле. Ошибки контраста связаны с наличием различного рода личностных (и не только личностных) установок. Феномен психологической проекции, выявленный психоанализом, также участвует в формировании этих установок.

§ 5. Проблемы, связанные с обработкой полученных данных

Основные "рецепты" по этому поводу достаточно просты. Главное — это не забывать, что, поскольку числовые данные представляют собой результаты измерений, то каждому уровню измерения (будь то шкала порядка или шкала интервалов) соответствуют определенные методы статистической обработки. Отметим основные моменты, на которые стоит обратить особое внимание:

1. При построении порядковой шкалы (как правило, метод балльных оценок для этого и используется) для усреднения повторных оценок одного испытуемого или при получении групповых баллов следует использовать не среднее арифметическое, а *медиану*. При обработке данных вручную для этого необходимо построить ранговый ряд и найти его середину. В качестве показателя вариативности полученных оценок используют не среднеквадратичное отклонение, а *межквартильный размах,* для чего необходимо построить частотное распределение исходных балльных оценок.

Как и в предыдущих заданиях, обработку целесообразно делать в статистической системе "Stadia". Для этого необходимо ввести исходные данные в электронную таблицу блока редактора данных, а затем войти в меню статистических методов (**F9**) и в нем выбрать первый пункт — "Описательная статистика". После нажатия на "Enter" и выполнения первых расчетов (среднее, дисперсия и т.д.) внизу экрана появится вопрос "Выдать дополнительную статистику?", на который нужно ответить утвердительно ("Y-да"), чтобы получить оценку медианы (Md) и квартилей (Q_1 и Q_3).

2. В том случае, если необходимо оценить корреляцию между двумя порядковыми (ранговыми) шкалами, правильным выбором будет использование непараметрического *коэффициента ранговой корреляции Спирмена*, а не коэффициента линейной корреляции Пирсона (как это часто делают). Последний адекватен лишь при измерениях не ниже шкалы интервалов. Для вычисления рангового коэффициента корреляции с помощью "Stadia" в меню статистических методов нужно найти раздел "Непараметрические методы" и выбрать в нем пункт "Корреляция (независимость)". После двух нажатий на "Enter" появляется значение коэффициента ранговой корреляции Спирмена — r и его статистическая значимость.

Литература

1. *Вудвортс Р., Шлосберг Г.* Психофизика II. Шкалирование. // Проблемы и методы психофизики / Под. ред. А.Г. Асмолова, М.Б. Михалевской. М.: Изд-во Моск. ун-та, 1974.

2. *Guilford J. P.* Psychometric Methods. N.-Y., Toronto, London: McGrow-Hill, 1954.

3. *Torgerson N.S.* Theory and Method of scaling. N.-Y.: John Wiley and Sons, 1958.

Методические указания по выполнению учебных заданий по теме "Метод балльных оценок"

В силу простоты подготовки стимульного материала к учебным заданиям по этой теме студентам предлагается самостоятельно подготовить и провести опыт с использованием метода балльных оценок для построения одномерной шкалы.

При планировании работы следует поставить определенную исследовательскую задачу. Например, сравнить эффективность использования двух различных процедур шкалирования, например, графического и числового методов или различных вариантов графического метода. Целесообразно, чтобы различные методы применялись на одном и том же материале — такое сравнение может наглядно показать методические преимущества одного из них.

Может возникнуть весьма интересная задача, если сравнить индивидуальную и групповую шкалы или две групповые шкалы, полученные на явно отличающихся выборках испытуемых. Исследование межгрупповых различий может быть очень интересным, если в качестве испытуемых взять людей из различных возрастных, социальных, религиозных или национальных групп и т.д.

Задачей исследования может быть сравнение шкал, построенных двумя различными методами одномерного шкалирования — методом балльной оценки и методом парных сравнений (см. следующую главу). В этом варианте, построив методом парных сравнений шкалу интервалов, будет весьма интересно сравнить ее со шкалой, полученной методом балльной оценки и проанализировать последнюю на предмет отражения в ней не только порядковых, но и, возможно, интервальных отношений между шкалируемыми объектами.

Один из обычных вариантов выполнения учебного задания с использованием метода балльных оценок — это его применение не как самостоятельного метода измерения, а в качестве вспомогательной процедуры получения балльных оценок при выполнении учебного задания по теме “Факторный анализ” (см. : ч. III, гл. 1 настоящего пособия). В этом случае выбор конкретной процедуры шкалирования будет определяться соответствующей содержательной задачей в контексте факторного анализа.

Те студенты, которые захотят выбрать такой компьютерный вариант выполнения задания, который требует сложно организованной и строго дозированной во времени стимуляции, могут воспользоваться специальной программой-конструктором психологических методик “StimMaker”. Эта программа позволяет достаточно просто и быстро в диалоге с компьютером спроектировать процедуру предъявления различных стимулов на экране монитора и регистрации ответных реакций испытуемого на каждый стимул. Данная программа позволяет создать на экране любого цвета стимул как комбинацию цифро-буквенных символов или символов псевдографики и задать любой порядок предъявления созданных стимулов. Такой вариант подготовки учебного задания вполне доступен всем, кто имеет хотя бы небольшой опыт работы на персональном компьютере.

В заключение в качестве примера мы приведем несколько вариантов учебных заданий, выполненных студентами факультета психологии Московского государственного университета им. В.М. Ломоносова в 1994—1995 г.г. :

1. Построение шкалы голосов знакомых и незнакомых людей.

2. Оценка популярности преподавателей факультета психологии среди студентов 2-го курса.

3. Оценка различных аспектов популярности учебных курсов, читаемых на факультете психологии во 2—3 семестрах.

4. Построение шкалы популярности современных рок-групп.

5. Построение шкалы популярности русских писателей второй половины XIX века.

6. Оценка популярности политических лидеров России в группах испытуемых так называемой “демократической” и ”леворадикальной” ориентации.

7. Построение шкалы “золотого” сечения: повторение эксперимента Г.Фехнера по оценке предпочтения прямоугольников.

Глава 2. МЕТОД ПАРНЫХ СРАВНЕНИЙ. МОДЕЛЬ ТЕРСТОУНА

§ 1. Закон сравнительных суждений

Самые распространенные в настоящее время методы шкалирования субъективных характеристик стимулов, не имеющих прямых физических коррелятов, основаны на модели шкалирования *Терстоуна* (Терстоун, 1927). Но первый шаг в этом направлении сделали *Фуллертон и Кэтелл* (1892), которые предложили подход, преобразующий постулат Фехнера о равенстве "едва заметных различий" в понятие равенства на континууме "равно часто замечаемых различий". Этот подход позволил перейти к оценке стимула, безотносительно к прямому физическому корреляту, но сразу же обнажилась проблема: если один стимул предпочитается второму с частотой *A*, а второй стимул предпочитается третьему с частотой в *1.2A*, то насколько субъективное расстояние между вторым и третьим стимулами больше субъективного расстояния между первым и вторым стимулами?

Торндайк (1910) предлагает решение этой проблемы (и это можно считать вторым шагом к цели), предположив, что разница в субъективных расстояниях пропорциональна различию в единицах стандартного отклонения нормальной кривой, соответствующих двум частотам.

Полное развитие этих идей и представляет собой модель шкалирования Терстоуна. Суть ее заключается в следующем:

1. Данное множество объектов можно *упорядочить в континуум* по какому-либо из параметров, который может служить стимулом, причем этот параметр не обязательно имеет физическую меру. Обозначим ряд стимулов как *1 ... i ... n*.

2. Каждый стимул теоретически вызывает у субъекта только один, свой *процесс различения* (обозначим его буквой S). Процессы различения составляют *психологический континуум, или континуум различения* (D_1 ... D_i ... D_n). Однако вслед-

ствие мгновенных флуктуаций организма, данный стимул может вызвать не только свой процесс различения, но и какие-то соседние. Поэтому, если один и тот же стимул предъявлять много раз, то на психологическом континууме ему будет соответствовать некоторое *распределение процессов различения*. При этом предполагается, что *форма распределения нормальна*.

3. В качестве значения i-го стимула на психологической шкале принимается *среднее* (S_i) распределения процессов различения, а дисперсия распределения рассматривается как *дисперсия различения* (σ_i).

4. Предъявление одновременно пары стимулов вызывает два процесса различения d_i и d_j. Разность (d_j - d_i) называется *различительной разностью*. При большом числе предъявлений двух стимулов различительные разности также формируют свое нормальное распределение на психологическом континууме. Поэтому среднее распределение разностей различения (d_j - d_i) будет равно разности средних распределений самих процессов различения — (S_j - S_i), а дисперсия распределения различительных разностей равна

$$s(d_j - d_i) = (s^2_j + s_i - 2r_{i,j} s_i s_j)^{1/2}, \qquad (1)$$

где s_i и s_j — дисперсии процессов различения i-го и j-го стимулов, соответственно, а $r_{i,j}$ — есть корреляция между мгновенными значениями процессов различения стимулов i и j.

Рассмотрим теперь следующую ситуацию. Пусть наблюдателю предъявляются пары стимулов i и j и от него требуется осуществить суждение, какой из стимулов дальше отстоит от нуля на психологическом континууме (например, более тяжелый или более сложный, или более красивый и т.д.). На рис. 1 показаны гипотетические процессы различения стимулов i и j.

Предполагается, что если различительный процесс для стимула j окажется на психологическом континууме выше, чем для стимула i, т.е. если различительная разность (d_j - d_i) > 0, то последует суждение, что стимул j больше, чем стимул i. И соответственно при (d_j - d_i) < 0 — произойдет обратное суждение.

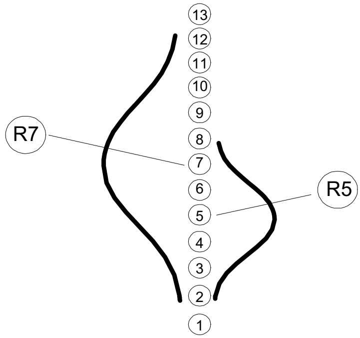


Рис.1. Гипотетическая модель процесса различения 2-х стимулов

Однако, если распределения различительных процессов перекрываются, то суждение, что стимул j меньше, чем стимул i может произойти даже тогда, когда величина S_j на психологическом континууме больше, чем величина S_i. На рис. 2 показано распределение различительных разностей при большом числе суждений.

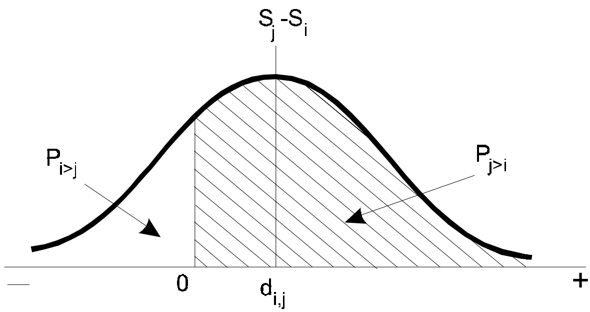


Рис. 2. Гипотетическое распределение процессов различения стимулов S_j и S_i на психологическом континууме: заштрихованная область указывает частоту суждения: стимул j больше, а незаштрихованная —я стимул j меньше; d_{ij} - различие шкальных значений стимулов i и j, измеренное в единицах стандартного отклонения данного распределения — $\sigma(d_j\text{-}d_i)$.

Среднее распределения равно различию шкальных величин двух стимулов — (S_j - S_i). Это различие можно найти из таблицы областей под единичной нормальной кривой, зная пропорцию суждений стимул j больше, чем стимул i от общего числа суждений по данной паре стимулов (т.е., сделав стандартное преобразование “$p \to z$”).

В единицах дисперсии $\sigma(d_j - d_i)$ это можно записать так:

$$S_j - S_i = z_{j,i}\sigma(d_j - d_i), \qquad (2)$$

где $z_{j,i}$ — обозначает искомое различие.

Подставляя это выражение в уравнение (1), получим:

$$S_j - S_i = z_{j,i}(\sigma_j^2 + \sigma_i^2 - 2r_{i,j}\sigma_i\sigma_j)^{1/2}. \qquad (3)$$

Уравнение (3) и выражает в общем виде *закон сравнительных оценок Терстоуна*.

§2. Процедура измерения

Эмпирическим материалом, на котором основан закон Терстоуна, служат суждения по типу: “стимул i более ... тяжелый, интересный, красивый и т.д., чем стимул j”. Прямой метод для получения таких оценок называется *методом парных сравнений*. В принципе это тот же самый метод константных стимулов, только в данном случае в качестве эталона выступает поочередно каждый стимул. Испытуемый осуществляет попарное сравнение всех стимулов. Каждое сравнение производится много раз. На основании этих сравнений для каждой пары определяется частота предпочтения одного стимула другому. Квадратная матрица (n x n) этих частот (обозначим ее буквой *F*) представляет исходные данные. Диагональные элементы этой матрицы будут пустыми, поскольку идентичные пары обычно не предъявляются. Очевидно, что сумма элементов $f_{i,j}$ и $f_{j,i}$ в сумме будет равна общему числу сравнений.

Последующий анализ заключается в переходе от матрицы частот (*F*) к матрице вероятностей (обозначим ее буквой *P*). Элемент этой матрицы $p_{i,j}$ есть пропорция числа предпочтений i-го стимула j-му в общем числе сравнений этих двух стимулов. Диагональ матрицы *P* также не заполнена, а сумма симметричных элементов относительно этой диагонали равна единице (т.е. $p_{i,j} + p_{j,i} = 1$). Из матрицы вероятностей уже легко определить матрицу различий *Z*, памятуя о том, что различие выражается в единицах нормального отклонения. Значение $z_{i,j}$ для соответствующей вероятности можно определить по таблице областей под единичной нормальной кривой. Для всех $p_{i,j}>0,5$ величина z будет положительна, а для всех $p_{i,j}<0,5$ — отрицательна. Для $p_{i,j}=1$ или $p_{i,j}=0$ $z_{i,j}$ не существует. Предполагая, что $p_{i,i} = p_{j,j} =0,5$, диагональные элементы матрицы *Z* приравниваются нулю. Поскольку $z_{i,,j} = -z_{j,i}$, то матрица будет косо-симметрична. Таким образом определяется матрица *Z*, элемент которой $z_{i,j}$ является оценкой различия (S_i - S_j) между шкальными значениями двух стимулов, измеренной в единицах стандартного отклонения в распределении различительных разностей. Каждый независимый элемент матрицы *Z* (а их, очевидно, будет n(n-1)/2) дает оценку различия для одного из уравнений (3) — как теоретической модели закона сравнительных оценок.

Рассмотрим теперь, как соотносятся исходные данные с теоретической формой их выражения. Число независимых элементов в матрице *F* равно n(n-1)/2, где n — число стимулов. Тогда как закон сравнительных оценок, выраженный в формуле (3), имеет для тех же n стимулов и n неизвестных шкальных значений, n неизвестных дисперсий различительных процессов и n(n-1)/2 неизвестных корреляций. Совершенно очевидно, что при таком соотношении числа уравнений — n(n-1)/2 и числа неизвестных — 2n+n(n-1)/2, решить данную систему невозможно. Поэтому необходимо ввести условия, упрощающие структуру выражения (3).

§ 3. Упрощенные варианты закона сравнительных суждений

Терстоун рассматривал 5 вариантов применения этого закона. Первый вариант — это та исходная общая форма закона, о которой уже говорилось. Второй вариант рассматривает изменение экспериментальной методики, обращаясь от оценок, производимых одним испытуемым, к групповым оценкам. Каждый испытуемый в этом случае производит только одно сравнение. И только третий, четвертый и пятый варианты вводят *дополнительные допущения*, которые меняют общую форму выражения (3).

Торгерсон (1958) предложил развести эти варианты на два класса. К первому классу относятся изменения в методике проведения эксперимента. Это первый и второй варианты Терстоуна, и кроме того, Торгерсон предложил отнести сюда и смешанный опыт, когда несколько испытуемых сравнивают по несколько пар и все оценки сводятся в общую матрицу частот. Ко второму классу относятся изменения в форме закона сравнительных оценок. Сюда относятся 3, 4 и 5 варианты Терстоуна или, соответственно, условия А, В и С, которые предложил Торгерсон.

III вариант Терстоуна. Предполагается, что корреляция между различительными процессами $r_{i,j}$ в выражении (3) равна нулю. В таком случае закон сравнительных оценок принимает форму:

$$S_j - S_i = z_{j,i}(\sigma_j^2 + \sigma_i^2)^{1/2}. \qquad (4)$$

Торгерсон предлагает здесь менее жесткое ограничение, с условием (условие А), что ковариация в выражении (3) — равна постоянной величине (d). Тогда :

$$S_j - S_i = z_{j,i}(\sigma_j^2 + \sigma_i^2 - d)^{1/2}. \qquad (5)$$

Но практически выражения (4) и (5) идентичны, поскольку ковариация является постоянной только тогда, когда корреляция стремится к нулю.

IV вариант Терстоуна основывается на допущении, что $r_{i,j}=0$ и что дисперсии различения мало отличаются друг от друга, т.е. $s_i = s_j + d$, где d мало по сравнению с s_j. Тогда выражение (3) преобразуется в

$$S_j - S_i = z_{j,i}[\sigma_j^2 + (\sigma_j + d^2)]^{1/2}. \qquad (6)$$

Раскрывая скобки и делая ряд преобразований и упрощений, получаем окончательную форму четвертого варианта закона:

$$S_j - S_i = z_{j,i}c(\sigma_j + \sigma_i), \qquad (7)$$

где c — постоянный множитель.

Более слабое допущение Торгерсона (условие B) о константности корреляции приводит к выражению:

$$S_j - S_i = z_{j,i}[1/2(1 - r)^{1/2} (\sigma_j + \sigma_i)]. \qquad (8)$$

Выражения (7) и (8) отличаются только постоянными членами, поэтому вариант Торгерсона имеет определенные преимущества.

<u>V вариант закона сравнительных оценок Терстоуна</u> нашел наибольшее применение вследствие простоты своей формы. Этот вариант основывается на допущении нулевой корреляции между двумя процессами различения ($r = 0$) и равенства различительных дисперсий этих процессов ($\sigma_j = \sigma_i = \sigma$). Тогда выражение (4) преобразуется в:

$$S_j - S_i = z_{j,i}\sigma \,. \qquad (9)$$

Обозначив константный член уравнения буквой “c”, получим:

$$S_j - S_i = cz_{j,i}. \qquad (10)$$

Уравнение (10) совпадает по своей общей форме с различными модификациями данного варианта, которые предлагали впоследствии некоторые авторы. Наиболее интересная модификация предложена Мостеллером (1951) и состоит в допущении равенства дисперсий и константной корреляции. В этом случае величина “c” в уравнении (10) будет равна $[2(1 - r)]^{1/2}$, а уравнение приобретает следующий вид:

$$S_j - S_i = z_{j,i}[2(1 - r)]^{1/2}. \qquad (11)$$

Сравнивая упрощенные варианты (4), (7), (10) с исходной формулой (3), легко видеть, что даже наиболее сложный из упрощенных вариантов (4) уже имеет, по крайней мере теоретически, решение, когда число стимулов (n) равно 5. Остальные варианты еще проще. Но практическая процедура всегда более трудоемка и менее изящна, чем это обещает теоретическая модель. Причина этого в основном лежит в эмпирической природе исходных оценок, в их зашумленности множеством случайных факторов, от которых невозможно оградить испытуемого. Для устранения случайных ошибок предлагается следующая тактика. Число стимулов увеличить так, чтобы система уравнений была значительно переопределена. Например, для варианта III брать не 5 стимулов, а 10 — 15. Для окончательного решения использовать итеративную вычислительную процедуру, которая учитывает тот факт, что случайные ошибки имеют тенденцию взаимоуравновешиваться.

Такие процедуры были разработаны разными авторами, и в данной работе будет описан алгоритм Мостеллера (1951) для V варианта закона в модификации Торгерсона (1958). Алгоритм использует решение методом наименьших квадратов. Он позволяет получить более точные оценки шкальных значений из матрицы в случае, если она не имеет пустых элементов.

§ 4. Процедура решения V варианта закона сравнительных оценок для полной матрицы

В V варианте закона, записанном в общем виде (9), единицы измерения шкальных значений всегда можно подобрать так, чтобы константа “с” была равна 1. Тогда:

$$S_j - S_i = z_{j,i}. \qquad (12)$$

В случае отсутствия ошибок в оценках искомое различие будет равно наблюдаемому (обозначим его $z'_{j,i}$). Но в результате ошибок между $z'_{j,i}$ и $z_{j,i}$ будет некоторое расхождение α. Задача заключается в получении такого множества оценок шкальных значений стимулов, для которых сумма квадратов всех расхождений является минимальной, т.е. необходимо минимизировать величину

$$\alpha_{i,j} = \sum_{i=1}^{n}\sum_{j=1}^{n} \left(z'_{i,j} - z_{i,j}\right)^2 \qquad (13)$$

Подставив вместо $z_{i,j}$ шкальные значения, получим:

$$\alpha_{i,j} = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(z'_{i,j} - S_i + S_j \right)^2 . \qquad (14)$$

Все $a_{i,j}$ для всех $z_{i,j}$ из матрицы Z дадут матрицу ошибок α. Чтобы минимизировать каждую $\alpha_{i,j}$, необходимо взять частную производную $\alpha_{i,j}$ по S_i и S_j. Каждое частное значение S_i в матрице ошибок a появляется только в i-той строке и i-том столбце, но поскольку матрица ошибок так же кососимметрична [$z_{i,j}$= -$z_{j,i}$ и (Si-Sj)= -(Sj-Si)], как и матрица Z, то для каждой S_i частная производная будет касаться только i-го столбца. Дифференцируя элементы каждого столбца по S_i, получим:

$$\frac{d\alpha_{i,j}}{dS_{i,j}} = -2\sum_{i=1}^{n} (z'_{j,i} - S_i + S_j) , \qquad (15)$$

где i = 1,2 ... , n.

Приравняем частную производную нулю и после переноса получим:

$$\sum_{j=1}^{n} z_{j,i} + \sum_{j=1}^{n} S_j = \sum_{j=1}^{n} S_i . \qquad (16)$$

Разделим выражение (16) на n и возьмем начальное значение шкалы, равное $\frac{1}{n}\sum_{j=1}^{n} S_i$. В результате получим:

$$S_i = \frac{1}{n}\sum_{j=1}^{n} z_{i,l} , \qquad (17)$$

где i=1,2 ... , n

Таким образом, для минимизации ошибки необходимо просто взять *среднее арифметическое по столбцу матрицы Z* и мы получим оптимальное значение шкальной величины S_i.

Рассмотрим практический пример решения V варианта закона сравнительных оценок методом наименьших квадратов (данные вымышлены). Испытуемому в случайном порядке предъявляются 6 цветных карт из малого набора теста Люшера и просят в каждой паре выбрать наиболее красивый. Каждая пара предъявляется по 50 раз. В итоге для одного из испытуемых была получена следующая матрица частот F (табл.1):

Таблица 1

Матрица частот — *F*

Стимулы	1	2	3	4	5	6
1	—	*29*	*35*	*42*	*46*	*49*
2	*21*	—	*26*	*33*	*42*	*45*
3	*15*	*24*	—	*26*	*32*	*43*
4	*8*	*17*	*24*	—	*28*	*34*
5	*4*	*8*	*18*	*22*	—	*28*
6	*1*	*5*	*7*	*16*	*22*	—

Примечание: элементом матрицы $f_{i,j}$ является частота, с которой в паре j,i стимул i оценивался более красивым, чем стимул j.

Полученная матрица частот F преобразуется в матрицу вероятностей P делением частоты $f_{i,j}$ на число предъявлений (N=50).

Таблица 2

Матрица вероятностей *P*

Стимулы	1	2	3	4	5	6
1	—	*0.56*	*0.70*	*0.84*	*0.92*	*0.98*
2	*0.42*	—	*0.52*	*0.66*	*0.94*	*0.90*
3	*0.30*	*0.48*	—	*0.52*	*0.64*	*0.86*
4	*0.16*	*0.34*	*0.48*	—	*0.58*	*0.68*
5	*0.08*	*0.16*	*0.36*	*0.44*	—	*0.56*
6	*0.02*	*0.10*	*0.14*	*0.32*	*0.44*	—
$\sum_{j=1}^{n} P_{j,i}$	*0.98*	*1.66*	*2.20*	*2.78*	*3.40*	*3.98*

Примечания: элементом матрицы $p_{i,j}$ является вероятность, с которой стимул i в паре j,i оценивался более красивым, чем стимул j.

Каждое значение вероятности $p_{i,j}$ из матрицы *P* переводится далее с помощью таблицы в единицы стандартного отклонения нормальной кривой — $z_{i,j}$, по которым и вычисляются шкальные значения S_i каждого стимула.

Таблица 3

Матрица *Z* - оценок

Стимулы	1	2	3	4	5	6
1	*0*	*0.20*	*0.52*	*0.99*	*1.42*	*2.05*
2	*-0.20*	*0*	*0.05*	*0.46*	*0.99*	*1.28*
3	*-0.52*	*-0.05*	*0*	*0.05*	*0.36*	*1.08*
4	*-0.99*	*-0.41*	*-0.05*	*0*	*0.15*	*0.47*
5	*-1.41*	*-0.99*	*-0.36*	*-0.15*	*0*	*0.47*
6	*-2.05*	*-1.28*	*-1.08*	*-0.47*	*-0.15*	*0*
$\sum_{j=1}^{n} z_{j,i}$	*-5.17*	*-2.53*	*-0.92*	*0.83*	*2.76*	*5.03*
$S_i = \frac{1}{n}\sum_{j=1}^{n} z_{j,i}$	*-0.86*	*-0.42*	*-0.15*	*0.14*	*0.46*	*0.84*

Примечания: элементом матрицы $z_{i,j}$ является вероятность $p_{j,i}$, преобразованная в единицы стандартного отклонения.

Рассмотренная процедура дает возможность для каждого стимула S_i получить его значение на *шкале интервалов.*

§5. Процедура решения V варианта закона сравнительных суждений для неполной матрицы исходных данных

Реальные экспериментальные данные очень часто отличаются от той классической матрицы данных, которая анализировалась выше. Наиболее распространенный *артефакт* в процедуре парного сравнения, который связан с ограничением на возможное число предъявлений, — стопроцентное предпочтение одного стимула другому, что приводит к появлению в матрице вероятностей нулей и единиц. Ноль и единица в терминах модели Терстоуна *не несут сравнительной информации* о различии стимулов, поэтому не могут быть использованы для расчетов шкальных значений стимулов.

Для матриц с нулями и единицами (они называются неполными матрицами) существуют особые алгоритмы анализа. Наиболее распространенный из них подробно описан в работе Торгерсона (1958) и вкратце состоит в следующем.

Из выражения (12) для стимула j следует, что стимул j+1 будет описываться следующим выражением:

$$S_{j+e} - S_i = z_{j,i} + e \ . \qquad (18)$$

Вычтя из уравнения (18) уравнение (12), мы получим сравнительное различие для интересующего нас стимула косвенным путем. В терминах минимизированной ошибки эта величина может быть вычислена из выражения:

$$d_{i,j+e} = S_{j+e} - S_j = \frac{1}{n_j}\sum_{i=1}^{n_j}(z_{i,j+e} + z_{i,j}) \ , \qquad (19)$$

где n_j — есть индекс суммирования.

Для практического удобства матрицу *Z* следует перестроить таким образом, чтобы столбцы были упорядочены по величине. Порядок столбцов в матрице *Z* определяется суммой по столбцу матрицы *P*. Для такой упорядоченной матрицы *Z* различие S_{j+e} - S_i можно прямо вычислить из выражения (19). Если мы шкальное значение первого стимула (S_i) приравняем к нулю, то шкальное значение любого стимула есть сумма шкального значения стимула и расстояния между данным стимулом и предшествующим:

$$S_1 = 0,$$

$$S_2 = d_{1,2},$$

$$S_3 = S_2 + d_{2,3},$$

$$S_n = S_{n-1} + d_{n-1,n}, \qquad (20)$$

Рассмотрим практический пример решения для неполной матрицы частот, взятый из работы Торгерсона (1958).

Пусть нам дана матрица вероятностей предпочтения i-го стимула j-му с некоторыми вырожденными (пустыми) элементами, равными 0 или 1.

Таблица 4

Матрица вероятностей *P*

Стимулы	1	2	3	4	5
1	—	*1.00*	*0.93*	*1.00*	*0.98*
2	*0.00*	—	*0.00*	*0.16*	*0.03*
3	*0.07*	*1.00*	—	*0.94*	*0.69*
4	*0.00*	*0.84*	*0.06*	—	*0.16*
5	*0.02*	*0.97*	*0.31*	*0.84*	—
$\sum_{j=1}^{n} P_{j,i}$	*0.09*	*3.81*	*1.30*	*2.94*	*1.86*

Примечания: элементом матрицы $p_{i,j}$ является вероятность, с которой стимул i в паре j,i оценивался более предпочтительным, чем стимул j.

Преобразуем вероятности $p_{i,j}$ в единицы стандартного отклонения нормального распределения — $z_{i,j}$.

Таблица 5

Матрица *Z — оценок*

Стимулы	1	2	3	4	5
1	*0*	—	*1.48*	—	*2.05*
2	—	*0*	—	*-0.99*	*-1.88*
3	*-1.48*	—	*0*	*1.55*	*0.50*
4	—	*0.99*	*-1.55*	*0*	*-0.99*
5	*-2.05*	*1.88*	*-0.50*	*0.99*	*0*
$\sum_{j=1}^{n} z_{j,i}$	*-3.53*	*2.87*	*-0.57*	*1.55*	*-0.32*

Примечания: элементом матрицы $Z_{i,j}$ является вероятность $p_{j,i}$, преобразованная в единицы стандартного отклонения.

Таблица 6

Матрица Z' — оценок

Стимулы	1	3	5	4	2
1	—	*1.48*	*2.05*	—	—
2	—	—	*-1.88*	*-0.99*	—
3	*-1.48*	—	*0.50*	*1.55*	—
4	—	*-1.55*	*-0.99*	—	*0.99*
5	*-2.05*	*-0.50*	—	*0.99*	*1.88*
$\sum_{j=1}^{n} z_{j,i}$	*-3.53*	*-0.57*	*-0.32*	*1.55*	*2.87*

Примечания: элементом матрицы $Z'_{i,j}$ *является вероятность* $p'_{j,i}$*, преобразованная в единицы стандартного отклонения. Столбцы упорядочены по возрастанию* $\sum p'_{j,i}$ *.*

Переставим столбцы в матрице Z в таком порядке, чтобы первый столбец имел наименьшую сумму элементов, а последний — наибольшую.

Таблица 7

Матрица разностей между столбцами

St / $d_{j,i}$	$d_{3,1}$	$d_{5,3}$	$d_{4,5}$	$d_{2,4}$
1	*1.48*	*0,57*	—	—
2	—	—	*0.89*	*0.99*
3	*1.48*	*0.50*	*1.05*	—
4	—	*0.56*	*0.99*	*0.99*
5	*1.55*	*0.50*	*0.99*	*0.89*
$\sum_{j=1}^{n} d_{j,i,i+1}$	*4.51*	*2.13*	*3.92*	*2.87*
Число эл-ов	*3*	*4*	*4*	*3*
$\frac{1}{n}\sum_{j}^{n} d_{ji,i+1}$	*1.50*	*0.53*	*0.98*	*0.96*

Из матрицы Z можно получить матрицу различий между соседними парами столбцов, вычитая их поэлементно один из другого. В каждой j-й строке элемент этой матрицы будет равен ($z_{j,i+1} - z_{j,i}$).

Пользуясь выражением (20), вычисляем из полученных различий шкальные значения стимулов, приняв, что $S_1 = 0$:

$S_1 = 0$,
$S_3 = 0 + 1.5 = 1.5$,
$S_5 = 1.5 + 0.53 = 2.03$,
$S_4 = 2.03 + 0.98 = 3.01$,
$S_2 = 3.01 + 0.56 = 3.97$.

Из рассмотренной процедуры видно, что недостающие элементы матрицы компенсируются наличием внутренней связи между элементами столбца, что позволяет рассматривать разность между столбцами матрицы как результат алгебраической интерполяции отсутствующих элементов в столбце.

Литература

1. *Терстуон Л.Л.* Психофизический анализ // Проблемы и методы психофизики / Под ред. А.Г.Асмолова, М.Б.Михалевской. М.: Изд-во Моск. ун-та, 1974.

2. *Guilford J. P.* Psychometric Methods. N. Y., Toronto, London: Mc-Grow-Hill, 1954.

3. *Torgerson N.S.* Theory and Method of scaling. N. Y.: John Wiley and Sons, 1958.

Методические указания по выполнению учебного задания по теме: “Метод парных сравнений”

Задание 1. Построение шкалы цветовых предпочтений методом парных сравнений

Цель задания: *Освоить метод парных сравнений для построения шкалы интервалов. Сравнить построенную шкалу со шкалой порядка, полученную методом балльной оценки.*

Методика

Аппаратура. Задание выполняется на IBM-совместимом персональном компьютере. Для предъявления сигнала “Внимание” используются головные телефоны, соединенные со звуковым синтезатором персонального компьютера. Для выполнения учебного задания используется компьютерная программа *parcom.exe* и *mbe.exe*[1].

Стимуляция. На экране монитора предъявляются цветные прямоугольники из набора восьмицветного теста цветовых предпочтений Люшера: синий, зеленый, красный, желтый, фиолетовый, коричневый, черный и серый.

Процедура опыта. При отработке задания каждый студент выступает сначала в роли испытуемого, а затем обрабатывает собственные данные. Испытуемый сидит на расстоянии 1 м от экрана дисплея. Опыт состоит из 2-х серий.

В первой серии испытуемому предлагается оценить по 10-балльной шкале *приятность* каждого цвета. Для этого на экране монитора ему предъявляется вертикальная графическая шкала с десятью оценочными градациями от “невообразимо приятный — 10 баллов” до “невообразимо неприятный — 0 баллов” . Внизу экрана в случайном порядке расположены 8 цветных прямоугольников. Используя клавиши управления движением курсора <←> и <→>, испытуемый может перемещать белую рамку от одного прямоугольника к другому и, таким образом, осуществлять свой выбор. Выбрав тот стимул, который нужно оценить, испытуемый нажимает на клавишу “Tab” и вводит нужное число от 0 до 10. Справа от графической шкалы на соответствующем месте появляется прямоугольник того же цвета, а в нижнем ряду он исчезает.

[1] Этот опыт можно проводить и без компьютера, имея набор стандартных цветовых карточек. Естественно, что в таком случае экспериментатор должен предварительно подготовить квазислучайную последовательность предъявления стимульных пар и протокол, куда будут заносится ответы испытуемого.

Действуя таким образом, испытуемый поочередно оценивает все 8 стимулов.

Во второй серии цветные прямоугольники предъявляются парами, и задача испытуемого заключается в том, чтобы оценить, какой из 2-х цветов ему нравится больше. Для ответа используются две клавиши управления движением курсора: <←> (левый нравится больше) и <→> (правый нравится больше). Как только испытуемый дает ответ, на экране появляется следующая пара стимулов. Всего предъявляются 144 пробы, т.е. все цвета встречается друг с другом по 6 раз. Три раза каждый из цветов предъявляется слева, три раза — справа. В верхнем правом углу экрана каждый раз высвечивается порядковый номер пробы.

Обработка результатов. После опыта студенту выдается компьютерная распечатка, в которой представлены: 1) по результатам первой серии — балльные оценки всех 8 цветов; 2) по результатам второй серии — усредненная по 6 предъявлениям матрица частот (F) — 8x8, элементом матрицы $f_{i,j}$ является частота, с которой в паре j,i стимул i оценивался более красивым, чем стимул j. При необходимости можно переписать на дискету файл с данными: его имя соответствует фамилии испытуемого, написанной латинскими буквами, а расширение — *mpc*.

Обработка результатов заключается в построении по каждой серии индивидуальной и групповой шкал[1]. По данным, полученным в первой серии, строится шкала порядка, по данным второй серии — шкала интервалов. Для получения групповых данных каждый испытуемый должен свести в таблицу и усреднить свои данные с данными других четырех испытуемых. Причем в академической группе студентов (как правило, 12 — 15 человек) не должно быть повторяющихся результатов.

Обсуждение результатов. При обсуждении полученных результатов каждый испытуемый должен сравнить располо-

[1] В случае получения по индивидуальным данным неполной матрицы вероятностей, т.е. состоящей из большого количества нулей и единиц, обрабатываются только групповые результаты. Будем считать матрицу неполной, если более 30% ее элементов, т.е. 23 и больше, равны нулю или единице.

жение стимулов по шкале порядка и шкале интервалов и сделать заключение о преимуществах и недостатках каждого метода. Стоит подумать о метрических преимуществах шкалы интервалов, и об отражении в шкальных значениях более тонких особенностей сходства или различия между стимулами. Кроме того, необходимо дать сравнительную оценку индивидуальной и групповой шкал.

Следует также сопоставить исходные положения модели с полученными в эксперименте результатами и сделать выводы (сравнительно с другими методами) о преимуществах и недостатках метода парных сравнений.

Глава 3. МЕТОДЫ ПРЯМОЙ ОЦЕНКИ

Согласно наиболее распространенной точке зрения группу методов, которые позволяют получить количественную оценку психологической величины на шкале интервалов или отношений как непосредственный результат измерительной процедуры, называют *методами прямого шкалирования*. Прямое шкалирование базируется на предположении о наличии у человека внутренней шкалы измеряемого психологического признака, имеющей начальную точку и единицу измерения, и, следовательно, способности выносить *количественные* суждения относительно своих ощущений. Другие, рассмотренные выше методы одномерного шкалирования (методы балльных оценок и парных сравнений), являются прямыми только для порядковой шкалы. Шкалы более высокого порядка могут быть построены при использовании этих методов только в случае введения дополнительных теоретических допущений, например, допущение о нормальном характере распределения измеряемого признака дает возможность получить вместо порядкового измерение на шкале интервалов.

Другой смысл названия прямого шкалирования состоит в том, что им подчеркивается *простота*, наиболее короткий путь от измерительной процедуры, в которой исследователь получает “сырые” данные, до построения субъективной шкалы, поскольку шкальное значение измеряемого психологического признака выражено в содержании ответа испытуемого. Конструирование шкал интервалов или отношений с помощью непрямых методов (косвенное шкалирование) требует помимо дополнительных теоретических допущений еще и ряда статистических манипуляций, т.е. осуществляется более опосредованным и трудоемким путем.

Во всех методах прямого шкалирования используются два типа организации ответных реакций испытуемого в ходе измерения: *процедура оценки*, когда от испытуемого требуется только сообщать свои суждения о предъявляемых ему объектах, и *процедура воспроизведения*, с помощью которой экспериментатор узнает о субъективных оценках испытуемого по тому, как он воспроизводит заданные величины, интервалы или ощущения, регулируя величину измеряемого параметра объекта.

К числу прямых методов, приводящих к интервальной шкале, *относится метод кажущихся равными интервалов* (другое его название — метод категориальной оценки), *метод последовательных интервалов и метод равных сенсорных расстояний* (иначе он называется методом воспроизведения категорий).

Шкала отношений может быть получена непосредственно из результатов измерения при использовании *метода оценки отношений, метода постоянных сумм, метода установления заданного отношения (фракционирование и мультипликация), метода оценки величины и метода воспроизведения заданной величины.*

Заслуга введения этих методов в широкую практику психологических измерений принадлежит С.С. Стивенсу. Ниже будут подробно рассмотрены два метода из перечисленных выше — метод установления заданного отношения и метод оценки величины.

§ 1. Метод установления заданного отношения

1. Модификации метода установления заданного отношения.

Известны две модификации метода установления заданного отношения: *метод фракционирования (деления) и метод мультипликации (умножения)*. В методе фракционирования (деления) испытуемому предъявляют поочередно несколько стандартных стимулов (S_{st}) и просят подобрать к каждому из них среди предъявляемых ему на сравнение стимулов (S_c) такие, величины которых составляют заданную часть от соответствующих S_{st}. Обычно задаются простые дроби типа $1/n = 1/2$, $1/3$ и т.п. Чаще всего используется $1/n = 1/2$, т.е. “деление пополам". При подборе стимула, находящегося в заданном отношении к S_{st}, используются процедуры оценки или воспроизведения.

Метод мультипликации (умножения) отличается от фракционирования только тем, что испытуемый должен подбирать к стандартному стимулу такой, который превышает его в заданное число раз, т.е. $n>1$.

2. Требования к шкалируемому признаку.

Методы фракционирования и мультипликации могут применяться для шкалирования только в том случае, когда шкалируемый признак удовлетворяет двум условиям:

1. Испытуемый должен иметь возможность наблюдать изменение переменного стимула, с помощью которого он подбирает заданное отношение, как непрерывное или с очень малым шагом. В противном случае он не сможет точно установить заданное отношение.

2. Субъективному шкалируемому признаку должна соответствовать физическая шкала стимулов, поскольку построение субъективной шкалы данным методом возможно только через определение психофизической функции, связывающей величины ощущений со стимульными значениями. Дело в том, что методом установления заданного отношения можно измерить (т.е. сопоставить со шкальными значениями одной и той же шкалы отношений) ощущения, вызванные не всеми, а только некоторыми стимулами (S_{st}). Например, если заданное отношение равно $1/n$, то однозначно могут быть определены шкальные значения ощущений, вызванных только стимулами, равными n^a, где $a = \pm 1,2,3,...,n$. Если $1/n = 1/2$, то в эксперименте можно получить только следующие шкальные значения: 1/16, 1/8, 1/4, 1/2, 1, 2, 4, 8, 16 ... Но для любого шкального значения, лежащего между соседними, вызванными стимулами n^a и n^{a+1}, в общем случае нельзя однозначно указать соответствующий ему стимул. Им может быть любой стимул, лежащий между n^a и n^{a+1}, и определить его можно только при наличии психофизической функции.

Процесс определения психофизической функции состоит из нескольких этапов:

1. Получение в процедуре измерения оценок стимулов, находящихся в заданном отношении к стандартным стимулам.

2. Проведение регрессионного анализа полученных данных, построение графика деления (умножения) на n и психофизической зависимости.

3. Проверка выполнения свойств шкалы отношений.

4. Определение вида психофизической зависимости.

Рассмотрим последовательно каждый из этих этапов на примере работы (Харпер, Стивенс, 1948) по взвешиванию грузов.

3. Организация измерительной процедуры.

При измерении оценок стимулов, находящихся в заданном к стандартным стимулам отношении, используется какой-либо из методов определения локализации точки на оси (методы границ, констант или подравнивания). Процедуры этих методов минимизируют возможность систематических ошибок в оценке. Выбранный психофизический метод задает порядок предъявления переменных стимулов (S_c), среди которых нужно выбрать находящийся в заданном отношении к стандартному, и способ вычисления средней оценки, поскольку всегда имеет место разброс оценок, получаемых в отдельных пробах и у разных испытуемых.

Сравниваемые стимулы выбираются так, чтобы охватить весь возможный диапазон разброса оценок при подборе стимула, находящегося в заданном отношении к стандарту, и обеспечить хорошую точность оценки, т.е. они изменяются малыми шагами. Они предъявляются либо рандомизированно (метод констант), либо восходящими и нисходящими рядами (метод подравнивания). Средняя оценка вычисляется как точка субъективного равенства (PSE). Число измерений на каждую PSE может быть уменьшено при хорошей “кучности” оценок. Обычно их число лежит в пределах от 30 до 100 на точку.

Стандартные стимулы выбираются так, чтобы охватить всю область измеряемого признака, а их число должно быть таково, чтобы обеспечить выявление разрывов психофизической функции, если они есть, и проведение гладкой кривой, если их нет. Как правило, используется не менее 5 стандартных стимулов. Обычно величины стандартных стимулов выбираются так, чтобы составить геометрический ряд, поскольку психофизическая зависимость чаще всего нелинейна.

Повторные оценки могут быть получены при опросе группы испытуемых, при повторном опросе одного испытуемого, а также обоими этими способами в зависимости от того, хотим мы получить эту шкалу для одного испытуемого или группы испытуемых. При повторном опросе одного испытуемого возникает вопрос, получать ли сразу несколько оценок для одного стандартного стимула, а затем переходить к следующему, или же получать одну оценку каждого из стандартных стимулов, а затем повторять всю серию. Наиболее предпочтительным с точки зрения независимости оценок является второй способ, который, однако, может оказаться более трудным для испытуемого.

Следует принять во внимание такой фактор, как тренированность испытуемых. Тренировка может уменьшить разброс оценок, т.е. увеличить их надежность, но вместе с тем процесс тренировки может изменить вид психофизической функции. Более того, различные способы тренировки могут привести к различным изменениям функции. Решение, тренировать ли испытуемых, зависит от того, как будет использоваться построенная шкала. Например, если в дальнейшем она будет применяться в работе с нетренированными испытуемыми, то не следует проводить их тренировки.

Предотвращение систематических ошибок и смещений, обусловленных внешними факторами. Причины смещений могут быть самыми разнообразными. Два хорошо известных примера — фиксированный временной или пространственный порядок предъявления переменного и стандартного стимулов приводит к появлению систематических смещений. Эти ошибки могут быть предотвращены посредством уравновешивающих процедур, предусмотренных в традиционных пороговых методах.

Несколько сложнее контролировать влияние так называемых контекстных эффектов. Многие исследования показали, что когда испытуемому предъявляют ряд переменных стимулов, он пытается выбрать как соответствующий заданному отношению со стандартом тот из стимулов, который расположен около середины ряда. Этот факт хорошо объясняется теорией уровня адаптации Хелсона. Влияние набора стимулов на суждение особенно сильно в тех случаях, когда оценка затруднительна для испытуемого. Гарнер (1954) показал, что выбор стимула, оцениваемого как половина стандарта, полностью зависит от используемого диапазона переменных стимулов. Гилфорд (1954) советует для полного устранения этого эффекта использовать один длинный ряд переменных стимулов для всех стандартных. Данные Стивенса и Поултона (1956) подтверждают, что контекстные эффекты исчезают, когда испытуемого не ограничивают фиксированным рядом сравниваемых стимулов, например, при использовании процедуры подравнивания.

Ниже приводится ряд стандартных стимулов весов, использовавшихся в работе Харпера и Стивенса (1948) и соответствующие им медианы (Md) весов, оцененных испытуемыми как равные половине стандартных (табл. 1).

Таблица 1

Результаты оценки испытуемыми стимула как половины стандартного (по Харперу и Стивенсу, 1948)

Вес стандартного стимула, г	Медиана оценок веса стимула, оцененного как 1/2 стандартного, г
20	15,8
40	28,0
70	51,7
100	77,0
300	195,0
500	337,0
1000	645,0
2000	1315,0

*4. Построение графика деления (умножения) на **n** и психофизической функции.* Средняя оценка стимулов, находящихся в заданном отношении *n* со стандартом, вычисляется либо как медиана Md, которая является грубой, но просто вычисляемой оценкой, либо как среднее геометрическое G, определяемое по формуле:

$$G = \sqrt{S'_1\, S'_2\, S'_3 \ldots S'_n}\ , \qquad (1)$$

где, S_1 ... S_n — величины стимулов, оцененных как составляющие заданную часть от стандартного; n — число повторных оценок .

Если число оценок больше трех, то G удобнее находить путем логарифмирования:

$$\lg G = \frac{\sum \lg S_i^{'}}{n}\ . \qquad (2)$$

Харпер и Стивенс воспользовались, как уже было сказано выше, медианой для оценки весов, воспринимаемых как половина стандартного. На основании полученных данных была определена зависимость $S' = f(S_{st})$, где S' — медиана стимулов, оцениваемых как половина стандартного стимула. Эта зависимость представлена на рис. 1.

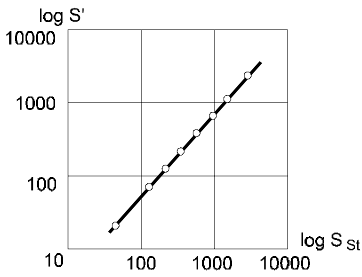


Рис. 1. График “деления на 2”:
по оси абсцисс — веса стандартного стимула (S_{st}), в граммах; по оси ординат — веса, воспринимаемые как половина от стандартных (S'), в граммах. Обе оси взяты в логарифмическом масштабе из-за большого диапазона значений стимулов. По экспериментальным точкам проведена регрессионная прямая (по Харперу и Стивенсу, 1948)

В данном случае экспериментальные точки почти точно ложатся на прямую, и она без явных ошибок может быть проведена на глазок.

Обычно линия, сопоставляющая на графике деления на *n* каждому стандартному стимулу S_{st} стимул S', воспринимаемый как объективно в *n* раз меньший, проводится через конечное и, как правило, небольшое число точек, соответствующих использованным стандартным стимулам. Проведение плавной линии через несколько точек, разумеется, всегда содержит ошибку, неточность. Однако, если вид зависимости известен (линейная, логарифмическая и т.п.), то неточность по отношению к экспериментальным точкам можно минимизировать. Минимизация ошибки является задачей регрессионного анализа, а полученная в результате решения этой задачи линия называется линией регрессии (прямолинейной, логарифмической и т.п.). Пока мы можем забыть о допускаемой неточности в определении этой кривой и рассматривать ее как непрерывную и "точную" для всех S.

Как от графика деления на *n*, который является только стимульно-стимульной функцией, перейти к психофизической функции? Для этого нужно только ввести *единицу измерения* на субъективной шкале, поскольку все нужные для построения субъективной шкалы соотношения субъективных и стимульных значений уже содержатся в полученной в опыте зависимости: $S' = f(S_{st})$. Для этого выбирается какой-либо из стандартных стимулов и соответствующее ему значение на шкале ощущения (Z) принимается за единицу (Z=1). Харпер и Стивенс выбрали в качестве такового ощущение тяжести, возникающее при поднятии груза 100 г., и назвали эту единицу "вег" (от старонорвежского слова, имеющего значение "поднимать"). Естественно, что шкальное значение того веса, который испытуемый оценил как вдвое менее тяжелый, чем S_{st} = 100 г, равно 1/2 вега. Это вес 77 г. В принципе метод установления заданного отношения позволяет указать любой стимул, которому соответствует шкальное значение, равное n^a, где a = 0, ±1, ±2 .... В нашем примере, где 1/n = 1/2, можно найти значения 1/4, 1/8, 1/16, 2, 4, 8, 16 и т.д. Как это делается показанно на рис. 2.

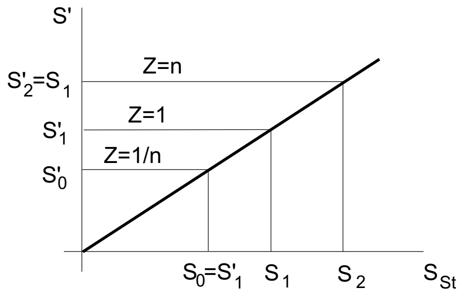


Рис. 2. Пример построения психофизической функции: по оси абсцисс — вес стандартного стимула, в граммах; по оси ординат — шкальные значения тяжести (Z).

Примем, что шкальное значение, соответствующее стимулу S_1, равно 1. Таким образом, мы вводим единицу измерения на будущей шкале (в нашем примере — это 1 "вег") и строим на ней первую точку с координатами (100; 1 или S_1; S'_1 на рис. 2). Тогда стимулу S_0, оцененному как в n раз меньший, соответствует шкальное значение 1/n. Отложив по оси абсцисс значение S_0 (мы его находим без труда из графика "деления на n", приведенного на рис.1, т.к. в опыте уже найден тот вес, который ощущается как половина от S_1), соотносим его со шкальным значением 1/n и строим на графике вторую точку. В нашем примере шкальное значение S'_0 будет равно 1/2. Так можно найти и все дальнейшие отрицательные степени n. Естественно, что точность построения психофизической функции будет зависеть от точности вычислений стимульных значений по графику "деления на n", что, в свою очередь, определяется "хорошестью" подгонки экспериментальных точек под плавную кривую или прямую, отражающую устойчивость полученной эмпирической зависимости. Чтобы получить все положительные степени того же отношения, необходимо изменить направление наших расчетов. Найдем по графику “деления на n” величину стимула, который при делении на n дает 1 вег — S'_2. Эту величину можно найти, проведя перпендикуляр от той точки на оси ординат, которая соответствует 1 вегу, до пресечения с аппроксимирующей кривой (прямой), и из точки пересечения опустить перпендикуляр на абсциссу. Найденная величина (S_2) соответствует n вегам (в нашем примере n = 2) и может, в свою очередь, быть использована для определения 2n вегов и т.д.

По найденным парам значений на субъективной шкале (Z) и на физической шкале стимулов (S) строится психофизическая функция: по оси абсцисс откладываются субъективные величины (например, веги), а по оси ординат — соответствующие им значения физического параметра стимула (например, граммы). Плавная линия, соединяющая точки, образованные парами значений Z и S, и образует графическую шкалу ощущений тяжести. Эта линия может быть проведена “на глазок” или с использованием методов регрессионного анализа и аппроксимирована подходящей математической функцией.

В дальнейшем психофизическая зависимость может использоваться для определения шкальных значений любого стимула, в том числе и такого, который не применяется в опыте, например, лежащего между S'_1 и S'_2. В самом деле, такому стимулу нельзя приписать однозначно шкальное значение, поскольку к нему нельзя “прийти” от предъявлявшихся в эксперименте стимулов S_1 или S_2 путем описанной выше процедуры с помощью кривой “деления на n”. Можно только утверждать, что его шкальное значение лежит между 1/n и 1. Это утверждение будет справедливо лишь при допущении, что психофизическая зависимость является строго монотонной. Неточность в определении шкального значения, соответствующего этому стимулу, возрастает за счет ошибки при построении психофизической зависимости.

Психофизическая функция, построенная по данным Харпера и Стивенса показана на рис. 3.

Аналитический способ, который дает более точное определение субъективной шкалы, поскольку лишен оши-

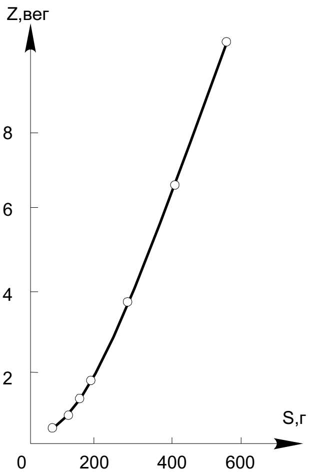


Рис. 3. Психофизическая функция тяжести поднимаемых грузов (Харпер и Стивенс, 1948)

бок, связанных с неточностью проведения графических работ, подробно описан Гилфордом (1954). Здесь приведем только краткую схему аналитического решения, поскольку для тех, кто владеет минимальными навыками регрессионного анализа, с помощью любой современной статистической программы оно не представляет большого труда[1]. Подобранные в опыте значения стимулов, оцененных как в *n* раз меньшие (большие), чем стандартные, преобразуются в логарифмы и с помощью метода наименьших квадратов определяется уравнение прямой. Качество подгонки полученной прямой под экспериментальные точки оценивается стандартным образом. Используя это уравнение, можно вычислить любое значение на оси "X" по известному значению на оси "Y" (и наоборот). Естественно, что точность получаемых оценок будет зависеть от качества полученной регрессионной прямой. Находя таким образом нужные значения на оси "X" конструируемой психофизической функции, получают все необходимые точки. После этого, применяя методы регрессионного анализа, определяют вид функции, описывающей психофизическую зависимость. Поскольку психофизические функции, как правило, нелинейны, удобнее представлять результаты на графике и проводить регресси-

[1] Конкретные методические рекомендации о том, как выполнить эту процедуру с помощью статистической системы "Stadia", будут даны ниже при описании учебного задания.

онный анализ в логарифмическом масштабе по оси абсцисс. Если эта функция подчиняется закону Фехнера, то в этом случае она будет прямой. Если же психофизическая функция степенная, то представление ее в виде прямой можно получить только в двойных логарифмических координатах (так называемые *log-log-координаты*), т.е. введя логарифмический масштаб также и по оси ординат. Таким образом, изображение психофизической функции в виде прямой в логарифмических координатах, является своеобразным "тестом" на ее соответствие одному из *основных психофизических законов*.

5. Проверка соответствия процедуры шкалирования шкале отношений: деление (умножение) на два взаимно простых числа.

Судя по приведенному выше описанию, метод фракционирования довольно груб с точки зрения получения точной психофизической зависимости. Оказывается, однако, что это не единственный и даже не самый главный его недостаток. Дело в том, что процедура этого метода не содержит возможности проверить, существует ли соответствие между выполненными испытуемым операциями отыскания стимула, относящегося как *1/n* к стандартному, и свойствами шкалы отношений. Следовательно, мы имеем повод сомневаться в том, действительно ли можно строить шкалу отношений по кривой деления (умножения) на *n*.

<u>Проверка выполнения свойств шкалы отношений.</u> Уточним, что следует понимать под "соответствием операций свойствам шкалы". В данном случае соответствие означает, что операция деления (умножения) стимула на число *n* (т.е. отыскания стимула, составляющего субъективно *1/n*-ю от стандарта) эквивалентна математической операции деления (умножения) наименованного числа (значения предполагаемой шкалы) на число-скаляр *n*. "Эквивалентна" означает, что она обладает теми же свойствами. Названная математическая операция обладает свойствами *ассоциативности, коммуникативности, тотальной сравнимости, обратимости и неизменности при умножении на 1.* Для наших целей достаточно представить эти свойства в виде следующих правил:

1. $Z = Z \cdot 1$ для любого шкального значения Z.

2. $Z \cdot \alpha_1 \cdot \alpha_2 \cdot \alpha_3 \; ... \cdot \alpha_n = Z \cdot \beta_1 \cdot \beta_2 \cdot \beta_3 \; ... \cdot \beta_n$, если и только если $\alpha_1 \cdot \alpha_2 \cdot \alpha_3 \; ... \cdot \alpha_n = b_1 \cdot b_2 \cdot b_3 \; ... \cdot b_n$ (это правило включает в себя и коммуникативность, и ассоциативность).

3. Для любых двух Z_1 и Z_2 существует единственное α, такое, что $Z_1 = Z_2 \cdot \alpha$ (тотальная сравнимость).

4. Если $Z_1 = Z_2 \cdot \alpha$, то $Z_2 = Z_1 \cdot 1/\alpha$ (это свойство обратимости)[1].

Рассмотрим, что означают эти правила на языке эмпирических операций деления (умножения):

1. Свойство 1 выполняется очевидно всегда, если только нет систематических ошибок, связанных с условиями эксперимента.

2. Пусть испытуемый “делит” стимул S на 2, тем самым он выбирает новый стимул S'_1. Стимул S'_1 он “делит” на 3 — выбирает стимул S'_2. Если бы первое “деление” было не на 2, а на 3, то вместо S'_1 должен был бы выбираться некоторый стимул S''_1. Правило 2 гарантирует, что если теперь S''_1 “разделить” на 2, то получится опять S'_2 (т.к. $1/3 \cdot 1/2 = 1/2 \cdot 1/3$). Этот пример, а также и другие примеры, демонстрирующие проверку правила 2, приведены на рис. 4.

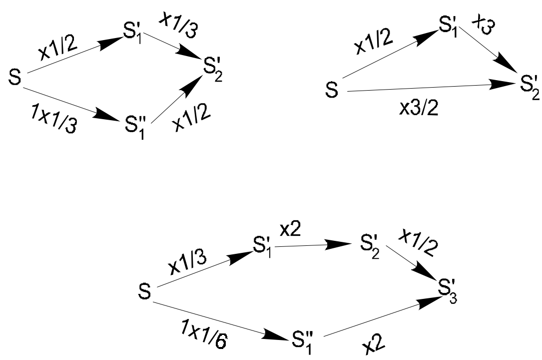


Рис. 4. Пример, демонстрирующий выполнение

---

[1] Это правило формально выводимо из правила 2, но для удобства его выделяют отдельно.

3. Правило 3 означает, что путем каких-то “умножений” и “делений” от одного стимула всегда можно “добраться” до любого другого. Если эксперимент организован так, что это правило выполняется, то мы избавляемся от необходимости строить психофизическую зависимость приблизительно (ведь до любого стимула можно “добраться” от “единичного” и тем самым получить точно соответствующее ему шкальное значение).

Можно доказать следующее утверждение: если экспериментально построены не одна кривая “деления на n” (см. рис. 1), а две — “деления на m” и “деления на n”, где n и m — взаимно простые числа (например, 2 и 3), то правило 3 выполняется. Доказательство следует из того факта, что любое шкальное значение может быть сколько угодно точно приближено числом вида $2^a \cdot 3^b (a,b = 0, \pm 1, \pm 2,...)$.

4. Правило 4 поясняется на рис. 5.

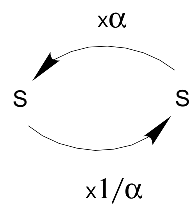


Рис. 5. Пример, демонстрирующий выполнение правила 4

Здесь, как и на рис. 4, стрелка обозначает выбор нового стимула. Проверка выполнимости правила может быть осуществлена так: строится кривая “деления на n” и кривая “умножения на n”, они должны совпасть с точностью до перемены осей (как функции *ln* и *exp*).

6. Определение вида психофизической зависимости.

Если бы возможный вид зависимости был совсем неизвестен, пришлось бы проделывать большую работу: провести регрессионный анализ для опытных данных, проверить выполнение свойств шкалы отношений, построить кривую психофизической зависимости и только после этого можно подбирать математическое выражение для полученной психофизической функции. Положение облегчается, если вид психофизической зависимости известен или по крайней мере должен быть осуществлен выбор между несколькими известными видами.

Известно, что большинство психофизических зависимостей может быть представлено в степенной или логарифмической форме. Рассмотрим основные варианты этих форм и те следствия, которые из них вытекают для кривых “деления” и “умножения”. Все эти следствия (хотя это и не будет доказываться) на самом деле являются не только необходимыми, но и достаточными условиями выполнения соответствующих форм психофизической зависимости.

1. Простейшая степенная форма $Z = aS^{\alpha}$. Какой вид должна иметь кривая “умножения” на n? Чтобы выяснить это, рассмотрим два значения стимула S и S_n, такие, что соответствующие им ощущения относятся как Z и Z · n:

$$Z = aS^{\alpha}, \tag{3}$$

$$Zn = a\,S^{\alpha}_{n}. \tag{4}$$

Разделим равенство (4) на (3):

$$n = \left(\frac{S_n}{S}\right)^{\alpha}; \qquad S_n = n^{1/\alpha} \cdot S. \tag{5}$$

Таким образом, если построить прямую наилучшего приближения по данным “умножения на n” (стимульно-стимульная кривая, где по оси абсцисс отложены значения стандартного стимула S, а по оси ординат — стимула, субъективно в n раз большего S', см. рис. 6), то:

1) прямая пройдет через начало координат (0,0);

2) наклон прямой покажет показатель степени в законе Стивенса. Этот показатель мы получим, если возьмем логарифм тангенса наклона (при основании, равном коэффициенту “умножения/деления” n), т.е. $\log_n \mathrm{tg}\varphi$, и вычислим обратную этому выражению величину (см. рис. 6).

2. Степенная форма $Z = k(S - S_0)^{\alpha}$ является степенной зависимостью с “порогом” (при $S = S_0$ ощущение равно 0, т.е. исчезает). Значения $S<S_0$ не рассматриваются. По аналогии с (3) и (4) запишем:

$$Z = k(S - S_0)^{\alpha}, \tag{3'}$$

$$Zn = k(S_n - S_0)^{\alpha}. \quad (4')$$

Разделив второе равенство на первое и проведя элементарные преобразования, получим:

$$S_n = n^{1/\alpha} S + (1 - n^{1/\alpha})S_0. \quad (5')$$

Итак, линия “умножения на n” оказывается прямой с наклоном n, но не проходит через начало координат (см. рис. 7).

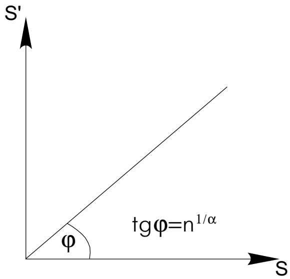


Рис. 6. Вид функции "умножение на n", проходящей через начало координат: по оси абсцисс — величина стандартного стимула; по оси ординат — величина стимула, оцененного в *n* раз больше, чем стандартный

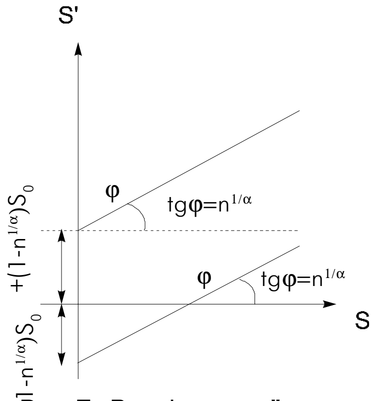


Рис. 7 . Вид функции "умножение на n", смещенной по оси ординат и не проходящей через начало координат:
по оси абсцисс — величина стандартного стимула, по оси ординат — величина стимула; оцененного в *n* раз больше, чем стандартный

Построив прямую наилучшего приближения по данным “умножения на n”, вычислим аналогично тому, как это делалось в предыдущем пункте А, показатель степени Стивенса. Однако, непрохождение прямой через (0,0) не позволяет ограничиться проделанным: недостаточно знать только α, нужно еще вычислить S_0. Прямая “умножения на n” пересекает ось ординат на уровне $(1 - n^{1/\alpha})S_0$. Разделив эту величину на $(1 - n^{1/\alpha})$, получим S_0.

На рис. 6 и 7 изображена прямая “умножения на n” в предположении, что $Z = aS^{\alpha}$ (рис. 6) и в предположении, что $Z = k(S_n - S_0)^{\alpha}$ (рис. 7). На рис.7 также показан случай, когда $(1 - n^{1/\alpha})S_0$ — величина отрицательная. Если S_0 действительно является “порогом”, то независимо от знака этой величины S_0 должна быть величиной положительной. Если этого не произойдет, то интерпретация S_0 меняется. Функция $Z = k(S+r)^{\alpha}$ (где $r>0$) показывает наличие “шума”, так что и при нулевом стимуле S_0 имеет место ненулевое ощущение $Z = kr^{\alpha}$. Эта разница в интерпретации не влияет на формальный анализ.

3. Простейшая логарифмическая зависимость $Z = logS$. В этом случае пара равенств, задающих кривую “умножения на n” такова:

$$Z = logS \, , \qquad (3')$$

$$Z = logS_n \, . \qquad (4')$$

Очевидно, что, проведя те же вычисления, как и в предыдущих пунктах, мы получим:

$$logS_n = nlogS, \qquad (5')$$

т.е. определенно нелинейную зависимость. Значит, если мы ожидаем логарифмическую, а не степенную зависимость, не следует строить прямых наилучшего приближения. Если

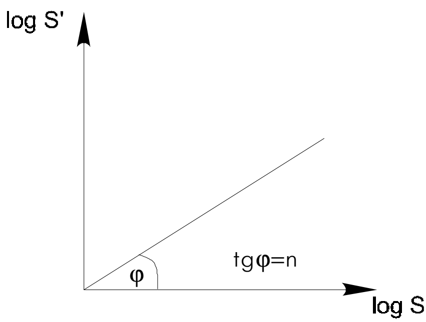


Рис. 8. Вид функции "умножение на *n*", проходящей через начало координат в двойных логарифмических координатах:
по оси абсцисс — логарифм стандартного стимула, по оси ординат — логарифм величины стимула, оцененного в *n* раз больше, чем стандартный

мы все же их построим, то они окажутся “плохими” в смысле приближения к опытным точкам, и самое главное, вычисления по разным n (n=1/2, 1/3 и 2) дадут нам разные величины a. Выход из затруднения состоит в том, что данные “умножения на n” следует откладывать в двойных логарифмических координатах. Тогда, согласно (2''), наилучшим приближением будет прямая, наклон которой равен коэффициенту фракционирования n (см. рис. 8).

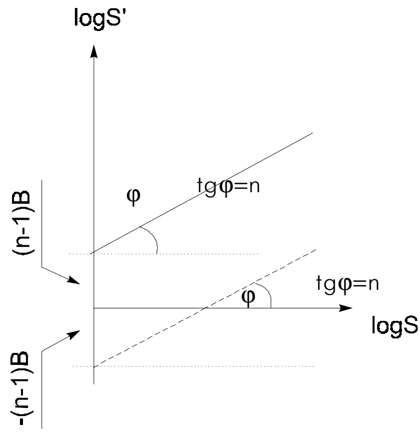


Рис. 9. Вид функции "умножение на n", смещенной по оси ординат и не проходящей через начало координат, в двойных логарифмических координатах:
по оси абсцисс — логарифм стандартного стимула; по оси ординат — логарифм величины стимула, оцененного в *n* раз больше, чем стандартный; штриховой линией показан случай, когда (n-1)b — величина отрицательная

4. Логарифмическая форма Z = logS + b. В этом случае имеем:

$$Z = logS + b, \qquad (3'')$$

$$Zn = logS_n + b. \qquad (4'')$$

Поделив второе равенство на первое и произведя элементарные преобразования, получим:

$$\log S_n = n \log S + (n-1)b. \qquad (5'')$$

График этой зависимости в двойных логарифмических координатах показан на рис. 9.

§2. Метод оценки величины

Метод оценки величины имеет своим предшественником метод дополнительного стимула, разработанный Меркелем еще в 1890 г., но потом прочно забытый. В современной форме метод оценки величины предложен *С. Стивенсом*. По его словам, "... все началось с дружеского спора с коллегой, который сказал: "Вы считаете, что у каждой громкости есть свое число и что, если кто-то издаст стон, то я смогу сообщить ему число, соответствующее этому стону". "Идея стоит того, чтоб ее испробовать", — ответил я. "Мы согласились, что как и в любой проблеме измерений сначала нужно решить вопрос о размере наших единиц. Я произнес громкий звук, обозначив его громкость как 100. Затем я предъявил ряд различных интенсивностей в случайном порядке и с готовностью, поразившей нас обоих, мой знакомый пронумеровал звуки в полностью сходной манере"[1]. Создавая этот метод, Стивенс стремился максимально *снять любые ограничения испытуемого в выражении своих впечатлений числом*, ограничения, связанные с введением обозначений концов стимульного ряда или с необходимостью определения отношений к заданному стандарту. Он хотел уменьшить какую бы то ни было предрасположенность испытуемого отвечать определенным образом в силу выбранной экспериментатором системы ответных реакций, например, отвечать только дробями.

Итак, *основное допущение* прямого шкалирования состоит в утверждении, что человек способен охарактеризовать числом величину любого своего впечатления, будь то приятность вкуса или громкость звука, красота произведе-

---

[1] Stevens S.S. The direct estimation of sensory magnitudes: Loudness// Amer. J. Psychol. 1956. Vd. 69, P. 1—25.

ния искусства или видимая яркость. Хотя прямое шкалирование применяется в основном в тех случаях, когда известен соответствующий измеряемый ощущениями физический континуум стимулов, по мнению Стивенса нет никаких принципиальных ограничений для прямого шкалирования и в тех случаях, когда исследователя интересует не психофизический закон.

Для получения шкалы методом оценки величины испытуемому должен быть предъявлен фиксированный ряд надпороговых стимулов, охватывающий достаточно широкий диапазон измеряемого признака. По утверждению Стивенса, средний испытуемый в оптимальных условиях способен оценить ощущения по шкале от 1 до 1000, вызванные стимулами, физическая интенсивность которых изменяется от 1 до биллиона (диапазон в 90 дБ). Как правило, в измерениях участвует много испытуемых ($n \geq 15$), но каждый дает мало оценок на каждый стимул (обычно всего 2). Стимулы предъявляются в случайном порядке. Довольно часто разным испытуемым предъявляются различные случайные последовательности стимулов. Действие временных факторов балансируется при получении второй оценки предъявлением стимульной последовательности в обратном порядке.

Существуют *2 формы метода* оценки величины: с заданным модулем или со свободным модулем (иначе его называют "без модуля").

1. Метод оценок величины с заданным модулем.

В начале опыта испытуемому предъявляется стандартный стимул и сообщается соответствующее вызванному им ощущению некоторое числовое значение на субъективной шкале признака — *модуль.* Все другие оценки испытуемый должен соотносить с этим модулем. Задача испытуемого подробно описывается в инструкции. В качестве примера взята инструкция из работы Энгена (1971) по шкалированию запахов:

“Мы хотим, чтобы Вы определили интенсивность запахов. Этот стимул представляет стандартную интенсивность. Другие запахи будут предъявляться в нерегулярном порядке. Все они примерно одинаковы по качеству, но их интенсивность различна. Назовем стандартный запах “10”. Ваша задача заключается в оценке интенсивности или силы всех других запахов относительно стандартного. Другими словами, если стандарт обозначен 10, как Вы обозначите сравниваемый запах? Используйте любые числа, которые кажутся Вам подходящими — дроби или целые. Например, если сравниваемый стимул пахнет в 7 раз сильнее стандартного, обозначьте его 70. Если он составляет 1/5 силы стандартного, оцените его 2, если 1/20, обозначьте 0.5 и т.д.

Здесь не может быть верного или неверного ответа. Мы хотим знать Ваше суждение об интенсивности запахов. Есть вопросы?”[1].

В ходе эксперимента стандарт предъявляется перед каждым переменным стимулом. Стивенсом (1956) были эмпирически выведены следующие правила организации опыта по оценке величин с модулем, позволяющие устранить влияния посторонних факторов на результат измерения:

1. Используйте стандарт, уровень которого не воспринимается наблюдателем как экстремальная точка в ряду интенсивностей, то есть используйте удобный стандарт, который, образно говоря, наблюдатель “мог бы подержать в руках”.

2. Предъявляйте такие переменные, которые были бы как выше, так и ниже стандарта.

3. Обозначайте стандарт числом, подобным 10, чтобы им легко было оперировать.

4. Назначьте число только стандарту и предоставьте наблюдателю полную свободу решать, как ему назвать переменный стимул. Например, не сообщайте наблюдателю, что самый слабый переменный стимул будет назван 1, или что самый громкий будет назван каким-нибудь другим числом. Если экспериментатор присвоит число более чем одному стимулу, он введет своего рода принуждение, которое заставляет наблюдателя производить суждения на более “сырой” — интервальной шкале, а не на шкале отношений.

---

[1] Engen T. Perception of odors. 1982. N. Y.: Academic Press, 1982.

5. Используйте только один уровень стандарта в серии, но в целом исследовании применяйте различные стандарты, так как рискованно судить о форме функции, полученной на основе данных только с одним стандартом.

6. Случайный порядок предъявления. С неопытным наблюдателем это хорошо, тем не менее начинайте с отношения громкости, которое не экстремально и которое, следовательно, легко оценить.

7. Делайте эксперимент коротким — около 10 мин.

8. Пусть наблюдатель сам себе предъявляет стимулы. Тогда он сможет работать своим темпом и это позволит ему быть более внимательным.

9. Так как оценки могут сильно отклоняться от оценок "среднего" наблюдателя, то желательно использовать группу наблюдателей, которая достаточно велика, чтобы получить при обработке устойчивую медиану.

В методе оценки величин шкальные значения измеряемого субъектом признака содержатся в ответах испытуемых и могут быть представлены медианой или геометрическим средним всех оценок группы испытуемых, относящихся к каждому из стимулов. Медиана является более грубой, чем геометрическое среднее, мерой центральной тенденции. Геометрическое среднее в отличие от среднего арифметического логически более обосновано, т.к. в результате измерения обычно получается нелинейная психофизическая функция типа $R = aY^n$. Геометрическое среднее определяется равенством (1). При большом числе измерений рекомендуется логарифмический вариант равенства (2).

Рассмотрим пример из упомянутой выше работы Энгена (1971), в которой каждый из испытуемых дважды оценивал 7 различных концентраций запаха амилацетата, разведенного в диэтилфтолате. В качестве модуля, которому приписывалось значение 10, предъявлялась концентрация 12.5^0. Геометрические средние оценок представлены в табл. 2.

Представление этих данных в линейных координатах показывает, что полученная психофизическая функция нелинейна и характеризуется замедляющимся ускорением: только при малых концентрациях амилацетата прирост запаха опережает рост концентрации. Будучи представлена в двойных логарифмических координатах, эта функция хорошо аппроксимируется прямой с наклоном <1. Следовательно, полученная психофизическая зависимость относится к числу степенных функций с показателем степени <1.

Таблица 2

Субъективная шкале запаха амилацетата, разведенного в диэтилфтолате

Шкальное значение	2.86	3.81	5.74	8.19	11.57	15.92	24.67
Стимул (концентрация)	1.56	3.12	6.25	12.50	25.00	50.00	100

Наиболее существенным недостатком метода оценки величин с заданным модулем может быть зависимость экспоненты степенной функции от места заданного модуля в стимульном ряду. Наличие такой зависимости (Энген, 1971) весьма неприятно тем, что ставит под сомнение экспоненту степенной функции как специальную характеристику модальности стимульного континуума. Однако существование такой зависимости далеко не всегда подтверждается экспериментами (Стивенс, 1956; Джонс и Восков, 1966).

2. Метод оценки величин со свободным модулем.

В этом варианте метода идея о независимости суждений испытуемого от выбора модуля получила логическое завершение: никакой стимул не объявляется стандартным, не вводятся никакие ограничения при выборе чисел для ответа. Единственное требование к испытуемому — стараться при ответе использовать числа, точно выражающие величину вызванного стимулом ощущения. Обычно стимулы предъявляются рандомизированно и в своем для каждого испытуемого порядке. Перед началом опыта испытуемому дают несколько (3—5) тренировочных проб.

Особенности этого варианта метода иллюстрируются в работе Кайн (1968) по оценке интенсивности запаха пентанола. 15 испытуемых оценивали каждую из 7 концентраций пентанола по 2 раза. Инструкция испытуемым:

"Вам будет предъявляться в нерегулярном порядке ряд тюбиков, содержащих один и тот же по качеству запах, но отличающихся по его интенсивности. Ваша задача — сообщать мне об интенсивности запаха, характеризуя его соответствующим числом. Когда Вы понюхаете тюбик, обозначьте интенсивность запаха числом — любым числом, которое Вам кажется подходящим. Затем я буду поочередно предъявлять Вам другие тюбики, и Вы будете каждому приписывать число. Постарайтесь, чтобы отношения между приписываемыми числами соответствовали отношениям между интенсивностями запахов. Иначе говоря, числа должны быть пропорциональны интенсивности исследуемого запаха. Помните, что вы можете использовать любое число, ограничений в выборе числа не существует. Правильного или неправильного ответа здесь быть не может. Нас интересует *Ваше* суждение об интенсивности запаха. Есть вопросы?"[1].

Полученные в экспериментах Кайн данные представлены в табл. 3.

Казалось бы, что геометрическое среднее каждой колонки матрицы может рассматриваться как значение на субъективной шкале силы запаха. Однако в силу отсутствия каких-либо ограничений в выборе модуля, используемого разными испытуемыми, числовые области могут сильно расходиться. Эта вариативность должна быть *устранена до* усреднения групповых данных.

Стивенс (1956) предлагает использовать простую процедуру приведения оценок каждого испытуемого, данных каждому из стимулов, к единой величине путем умножения на подходящий коэффициент. Процедура обработки "сырых" данных включает следующие этапы: 1) определение медианы или геометрического среднего двух оценок каждо-

[1] Цит. по : Engen T. Perception of odors. N. Y.: Academic Press, 1982.

го стимула каждым испытуемым; 2) выбор единого шкального значения оценки стимула (желательно брать его для середины стимульного ряда) и приведение всех оценок каждого испытуемого к единому масштабу через умножение на соответствующий коэффициент; 3) вычисление геометрического среднего или медианы для каждого стимула по приведенным оценкам всех испытуемых даст шкальные значения измеряемого признака.

Таблица 3

Индивидуальные оценки концентраций пентанола (Кайн, 1968)

Испытуемый	Концентрация пентанола, %						
	100	50	25	12.50	6.25	3.12	1.56
1	20.0	10.0	4.0	10.0	2.5	1.0	0.5
1	50.0	7.5	20.0	7.5	5.0	1.5	0.5
2	10.0	5.0	5.0	1.0	2.0	3.0	0.5
2	8.0	3.0	5.0	1.0	2.0	2.0	2.0
3	30.0	25.0	10.0	5.0	10.0	2.0	5.0
3	50.0	30.0	20.0	5.5	1.0	10.0	1.0
4	140.0	30.0	70.0	85.0	15.0	20.0	8.0
4	70.0	45.0	50.0	10.0	3.0	2.0	3.0
5	20.0	12.0	8.0	2.0	1.0	1.0	1.0
5	25.0	10.0	15.0	4.0	3.0	1.0	2.0
6	20.0	15.0	10.0	12.0	12.0	10.0	7.0
6	30.0	25.0	10.0	25.0	17.0	7.0	5.0
7	25.0	13.0	10.0	10.0	3.0	6.0	3.0
7	20.0	18.0	15.0	4.0	8.0	7.0	5.0
8	40.0	20.0	20.0	15.0	12.0	7.0	10.0
8	60.0	25.0	18.0	10.0	15.0	5.0	15.0
9	30.0	25.0	28.0	15.0	10.0	4.0	2.0
9	35.0	35.0	40.0	25.0	18.0	3.0	1.0
10	10.0	5.0	3.3	5.0	10.0	1.0	1.0
10	7.5	2.0	3.3	2.0	2.0	1.0	1.0
11	150.0	200.0	250.0	100.0	50.0	25.0	10.0
11	125.0	150.0	100.0	80.0	130.0	90.0	75.0
12	30.0	12.0	7.0	15.0	12.0	9.0	2.0
12	25.0	10.0	10.0	5.0	5.0	7.0	10.0
13	80.0	65.0	10.0	20.0	10.0	5.0	1.0
13	85.0	20.0	15.0	10.0	5.0	4.8	1.0
14	20.0	16.0	10.0	6.0	30.0	7.0	2.0
14	18.0	15.0	8.0	5.0	90.0	3.0	1.0
15	60.0	50.0	45.0	40.0	20.0	15.0	15.0
15	85.0	45.0	25.0	30.0	25.0	10.0	10.0

Достоинством предлагаемого Стивенсом способа обработки является его простота. Недостаток этого способа состоит в том, что определение коэффициента для приведения индивидуальных оценок осуществляется на основе суждений только об одном стимуле. Следовательно, в этом способе не учитывается индивидуальная несистематичность в использовании чисел, в силу которой оценка испытуемым отдельного стимула может заметно выпадать из общей закономерности. Если коэффициент приведения вычисляется по оценке именно такого стимула, то возрастает ошибка в усреднении групповых оценок всех стимулов.

Энген (1971) предлагает более громоздкий, но зато и более корректный способ первичной обработки экспериментальных данных, в котором учитывается как меж-, так и внутрииндивидуальная вариативность суждений. Его способ состоит из 6 этапов:

1. Определить логарифм каждой численной оценки стимула.

2. Вычислить логарифм геометрического среднего оценок каждого стимула в отдельности каждым испытуемым.

3. Найти логарифм геометрического среднего оценок каждым испытуемым всех стимулов, найти среднее каждого ряда (см. табл. 3).

4. Определить среднее всех величин, вычисленных на 3-м этапе — логарифм общего геометрического среднего оценок всех стимулов всеми испытуемыми в матрице “сырых” данных.

5. Вычесть величину логарифма общего геометрического среднего (п.4) из логарифма индивидуального геометрического среднего (п.3).

6. Сложить разности, полученные в п.5, с соответствующими значениями логарифмов геометрического среднего оценок испытуемым каждого стимула (п.2).

Окончательные результаты такой первичной обработки данных, приведенных в табл. 3, показаны в табл. 4.

Цель, которая достигается этим способом обработки, состоит в уменьшении разброса индивидуальных данных вокруг основной функции, но не влияет на ее параметры.

Для получения шкальных значений нужно вычислить среднее для каждой колонки, антилогарифм которого и является геометрическим средним оценок группой испытуемых данного стимула. Найденные таким образом геометрические средние используются при определении вида психофизической зависимости.

Таблица 4

Скорректированные оценки испытуемых, представленные в табл. 3 (по Энгену)

Испытуемый	Концентрация пентанола, %						
	100	50	25	12.50	6.25	3.12	1.56
1	1.84	1.27	1.29	1.27	0.89	0.43	0.04
2	1.54	1.17	1.28	0.59	0.89	0.97	0.59
3	1.68	1.53	1.24	0.79	0.59	0.74	0.44
4	1.70	1.27	1.47	1.17	0.53	0.50	0.39
5	1.74	1.43	1.43	0.85	0.63	0.39	0.54
6	1.28	1.18	0.89	1.13	1.04	0.82	0.66
7	1.42	1.26	1.16	0.87	0.76	0.88	0.66
8	1.50	1.15	1.08	0.89	0.93	0.58	0.89
9	1.43	1.39	1.44	1.20	1.04	0.46	0.07
10	1.42	0.98	1.00	0.98	1.13	0.68	0.83
11	1.20	1.31	1.26	1.02	0.97	0.74	0.50
12	1.47	1.07	0.96	0.97	0.92	0.94	0.69
13	1.78	1.70	1.38	1.02	0.72	0.56	0.13
14	1.47	1.38	1.14	0.93	0.91	0.85	0.34
15	1.41	1.23	1.08	1.10	0.91	0.65	0.65
Σlog	22.68	19.32	18.10	14.78	12.86	10.19	7.15
Средн. log	1.53	1.29	1.21	0.99	0.86	0.68	0.47
Геом. ср.	33.5	19.4	16.2	9.7	7.2	4.8	3.0

В том случае, если эта зависимость подчиняется степенному закону $Z = kS^{\alpha}$ при изображении ее в двойных логарифмических координатах, она должна быть прямой линией. Если же исследуемая зависимость подчиняется закону Фехнера Z = klogS, она может быть представлена прямой только в том случае, если логарифмический масштаб берется только для шкалы физического параметра стимулов, а шкала суждений остается линейной.

Как было отмечено выше, прямая удобнее (проще) с точки зрения наилучшей подгонки к экспериментальным точкам и подбора математического описания полученной зависимости. Обычно подгонка осуществляется либо “на глазок”, что дает грубое приближение при разбросе точек вокруг истинной линии регрессии, либо методом наименьших квадратов, позволяющим обеспечить наилучшую аппроксимацию, если заранее известен вид зависимости.

Основная

Литература

1. *Джелдард Ф.А.* Сенсорные шкалы//Хрестоматия по ощущению и восприятию/Под ред. Ю.Б. Гиппенрейтер, М.Б. Михалевской. М.: Изд-во Моск. ун-та, с. 1975.

2. *Стивенс С.С.* Психофизика сенсорной функции//Хрестоматия по ощущению и восприятию/Под ред. Ю.Б. Гиппенрейтер, М.Б. Михалевской. М.: Изд-во Моск. ун-та, 1975.

Дополнительная

1. *Стивенс С.С.* О психофизическом законе//Проблемы и методы психофизики/Под ред. А.Г. Асмолова, М.Б. Михалевской. М.: Изд-во Моск. ун-та, 1974.

2. *Harper R.S., Stevens S.S.* A Psychological Scale of Weight and a Formula for its Derivation. Amer. J. Psychol. 1948. Vd. 61. P. 343—351.

3. *Stevens S.S.* The Direct Estimation of Sensory Magnitudes: Loudness//Amer. J. Psychol. 1956. Vd. 69. P. 1—25.

Методические указания по выполнению учебного задания по теме: “Методы прямой оценки”

Задание 1. Построение психофизической функции высоты звукового тона

Цель задания. *Практически ознакомиться с методом установления заданного отношения, построить психофизическую функцию высоты звукового тона.*

Методика

Аппаратура. Задание выполняется на IBM-совместимом персональном компьютере. Для предъявления звуковых стимулов используются специальная звуковая плата типа “Sound Blaster” и головные телефоны. Для выполнения учебного задания используется компьютерная программа *sscal.exe*.

Стимуляция. Стимулами являются бинаурально предъявляемые тональные звуковые сигналы в диапазоне от 125 до 12000 Гц и интенсивностью около 80 дБ (УЗД). Используются 8 стандартных стимулов: 125, 250, 500, 1000, 2000, 4000, 6000, 8000 Гц.

Испытуемые. Каждый студент сначала работает как испытуемый, а затем обрабатывает свои оценки и строит по ним психофизическую зависимость. Для получения групповых данных испытуемые должны объединиться в группу не менее 7—10 человек.

Процедура. Шкалирование производится методом оценки заданного отношения в вариантах фракционирования (деления на 2) и мультипликации (умножения на 2). Опыт состоит из 2-х серий — “деление на 2” и “умножение на 2”. В первой серии используются 7 стандартных стимулов: частотой 250, 500, 1000, 2000, 4000, 6000 и 8000 Гц. Во второй — 125, 250, 500, 1000, 2000, 4000, 6000 Гц. Стандартные стимулы предъявляются в случайном порядке, каждый по 5 раз. Оценка заданного отношения производится методом установки, когда, регулируя частоту переменного стимула, испытуемый подбирает подходящий переменный стимул.

Каждая проба начинается с предъявления на экране монитора ее порядкового номера (от 1 до 35). После этого, нажав на клавишу <пробел> на клавиатуре компьютера, испытуемый должен прослушать звучание стандартного стимула. Для прекращения предъявления стандартного стимула следует нажать клавишу <Esc>. Регулировка частоты переменного стимула (от частоты стандарта) осуществляется клавишами управления движением курсора — <←> , <→> , для уменьшения и увеличения частоты, соответственно. Нажимая на эти клавиши, испытуемый может изменять частоту звукового сигнала: чем дольше нажимаешь на клавишу, тем быстрее меняется его частота. В любой момент испытуемый может вернуться к прослушиванию стандартного стимула, нажав еще раз на <пробел>. Окончание подбора переменного стимула, находящегося в заданном отношении к стандартному, подтверждается нажатием на клавишу <Enter>.

Между сериями устраивается 5-минутный перерыв.

Обработка результатов. После окончания опыта испытуемому выдается распечатка индивидуальных результатов, в которой для каждой серии (деление и умножение на 2) приводятся геометрические средние его оценок стандартных стимулов.

В ходе дальнейшей обработки индивидуальных результатов каждому студенту следует выполнить следующее:

1. Занести полученные данные в таблицу.

2. Представить полученные данные в линейных координатах на 2-х отдельных графиках (рис.1 и 2): абсцисса — частота стандартного стимула, ордината — частота стимула, оцененного находящимся в заданном отношении (1/2, 2) к стандартному.

4. То же самое повторить в двойных логарифмических координатах (рис. 3 и 4). С помощью линейного регрессионного анализа вычислить параметры регрессионных прямых для двух серий опыта и построить соответствующие прямые на рис. 3 и 4.

Для выполнения регрессионного анализа целесообразно воспользоваться статистическим пакетом “Stadia”. В редакторе данных частоты стандартного стимула заносятся в первую переменную (это будут значения X), а оценки испытуемого — во вторую (это будут значения Y). Затем переходят в меню статистических процедур (F9) и выбирают опцию “Простая регрессия (тренд)”. Войдя в нее, нужно указать номера переменных (1, 2), а затем указать тип функции для построения регрессии — линейная. После этого программа построит для вас математическую модель ваших данных, представляя их в виде уравнения прямой: $Y = a_0 + a_1*X$. Получив коэффициенты a_0 и a_1, можно без труда построить на графике аппроксимирующую прямую, проходящую через 7 экспериментальных точек. Статистическая оценка адекватности сделанной линейной аппроксимации приводится внизу экрана.

5. Построить психофизическую зависимость, используя результаты 2-х серий опыта — “деление на 2” и “умножение на 2”. Для этого поступим аналогично С. Стивенсу и соавт. (1937) и примем ощущение высоты тона 1000 Гц за точку отсчета, присвоив ей величину 1000 “мел”. Для нахождения необходимых значений на оси абсцисс психофизической функции можно воспользоваться либо графиками, построенными на рис. 3 и 4, либо решить ту же задачу более точно — аналитически.

С этой целью можно обратиться к уравнениям регрессионных прямых и проделать все вычисления по формуле вручную или еще раз воспользоваться программой “Stadia”. Формально это означает, что по полученному уравнению регрессионной прямой (геометрической модели оценок испытуемого) нужно найти неизвестные X по известным Y. Для проведения расчетов нужно снова вернуться в меню статистических методов и, выбрав ту же опцию (“Линейная регрессия”), указать другой порядок переменных — 2,1. Это будет означать , что в качестве X мы выбираем полученные оценки испытуемого, а в качестве Y — частоту стандарта. После расчета нового регрессионного уравнения можно найти все необходимые точки на оси X конструируемой психофизической функции.

Нанеся 8—10 точек на график психофизической функции можно попробовать соединить их плавной кривой и прикинуть “на глазок” вид полученной функции. Для более строгой оценки полученной зависимости следует вновь обратиться к статистической программе “Stadia” и, воспользовавшись методами регрессионного анализа, оценить ее как соответствующую логарифмической или степенной функции. Степень соответствия можно оценить по максимуму коэффициента корреляции и минимуму стандартного отклонения и доверительного интервала коэффициентов a_0 и a_1.

6. Построить психофизическую функцию по групповым данным, для чего вычислить геометрические средние по 7—10 индивидуальным оценкам, причем в каждую группу не должны входить одинаковые индивидуальные данные.

Обсуждение результатов.

1. Сопоставив данные умножения и деления, оценить обоснованность полученной субъективной шкалы, как шкалы отношений.

2. Сделать вывод о приложимости одного из известных психофизических законов к полученным экспериментальным данным.

3. Сравнить индивидуальную и групповую шкалы и оценить их адекватность литературным данным (см. статью Р. Вудвортса и Г. Шлосберга в кн. “Проблемы и методы психофизики”. Под ред. А.Г. Асмолова, М.Б. Михалевской, с. 183-186).

ЧАСТЬ III
МНОГОМЕРНОЕ ШКАЛИРОВАНИЕ

Глава 1. ФАКТОРНЫЙ АНАЛИЗ

Введение

Факторный анализ (ФА) принадлежит к числу таких методов, которые будучи разработанными в рамках запросов одной науки, впоследствии приобрели более широкое междисциплинарное значение. Заслугой психологии можно считать разработку именно такого метода. Исторически возникший в лоне психометрики, ФА в настоящее время приобрел статус общенаучного метода и широко распространен не только в психологии, но и в нейрофизиологии, социологии, политологии, экономике и статистике. Поэтому не стоит удивляться, если на вопрос к математику, что такое ФА, вы получите ответ, что этот метод понижения размерности корреляционной матрицы, а экономист добавит, что ФА используется им как средство визуализации многопараметрического объекта экономического анализа.

ФА как общенаучный метод, получивший к настоящему времени солидное математико-статистическое обоснование, имеет непростую историю, начиная с полного отказа математиков признать ценность используемого психологами в известной степени нестрогого и зависящего от мастерства исследователя алгоритма до обязательного включения различных вариантов ФА в любую сколько-нибудь известную компьютерную статистическую программу. Основные идеи ФА (впрочем, как и других методов многомерного анализа данных) были заложены в трудах известного английского психолога и антрополога *Ф. Гальтона* (1822—1911), основателя *евгеники* и внесшего также большой вклад в исследование индивидуальных различий. В разработку ФА внесли вклад многие ученые и только на русском языке опубликовано более 10 монографий, посвященных непосредственно ФА, однако психологи должны быть особенно признательны трем своим коллегам, с именами которых связаны разработка и внедрение ФА в психологию — это *Ч. Спирмен* (1904, 1927, 1946), *Л.Терстоун* (1935, 1947, 1951) и *Р. Кеттел* (1946, 1947, 1951). Кроме того, нельзя не упомянуть еще трех ученых — английского математика и философа *К. Пирсона*, в значительной степени развившего идеи Ф. Гальтона, американского математика *Г. Хоттелинга*, разработавшего современный вариант *метода главных компонент*, и известного английского психолога *Г. Айзенка*, широко использовавшего ФА для разработки психологической теории личности.

Необходимо отметить, что в силу профессиональных установок авторов в литературе на русском языке (К. Иберла, 1980; Г. Харман, 1972) ФА чаще всего излагается как один из методов математической статистики, а ориентированное на психологов изложение — скорее исключение, чем правило (Дж. Ким, Ч. Мьюллер, 1989; Я. Окунь, 1972). В данной главе ФА будет излагаться как один из методов психологического шкалирования и многомерного анализа данных. Кроме того, в отличие от других авторов, в силу ряда причин описывавших преимущественно *центроидный метод* ФА, мы уделим особое внимание более современным и компьютеризованным процедурам применения ФА. По возможности мы будем исключать экскурсы в математические основы той или иной процедуры, а сосредоточимся на описании основных этапов работы с эмпирическими данными и их трансформацией в ходе ФА. В силу специфики курса “Психологические измерения” в Общем психологическом практикуме изложение материала будет сопровождаться иллюстрациями использования 2-х статистических программ — “Stadia” и SPSS.

§ 1. Область применения факторного анализа

Необходимость применения ФА в психологии как одного из методов многомерного количественного описания (измерения, анализа) наблюдаемых переменных в первую очередь следует из многомерности объектов, изучаемых нашей наукой. Сразу же определим, что под многомерным представлением объекта мы будем понимать результат его оценивания по нескольким различным и существенным для его описания характеристикам-измерениям, т.е. присвоение ему сразу нескольких числовых значений. Из этого вполне естественно следуют два очень важных вопроса: насколько существенны и различны эти используемые характеристики. Первый вопрос связан с всесторонностью и полнотой описания объекта психологического измерения, а второй (в большей степени) — с выбором некоторого минимально разумного количества этих характеристик. Поясним сказанное выше на примере. Чем отличается хороший, профессионально сделанный психологический опросник от многочисленных "полупродуктов-полушуток", во множестве публикуемых в периодической печати для широкой публики или в книгах непрофессионалов-дилетантов? Прежде всего тем, что в первом случае объект психологического измерения (конструкт) описывается разносторонне и полно, и, кроме того, в нем не содержится лишних, не относящихся к делу (т.е. "не работающих" на ту или иную шкалу) вопросов. Таким образом, при использовании методов многомерных измерений психологических характеристик наиболее важны две проблемы — описать объект измерения *всесторонне* и, в тоже время, *компактно.* С известной долей обобщения можно сказать, что это одни из основных задач, решаемых ФА.

Понятно, что информативность многомерного описания объекта нашего изучения возрастает с увеличением количества используемых признаков или измерительных шкал. Однако очень трудно выбрать сразу и существенные, и независимые друг от друга характеристики. Этот выбор порой непрост и долог. Как правило, исследователь начинает с заведомо *избыточного* количества признаков, и в процессе работы (например, по созданию нового опросника или анализу экспериментальных данных) сталкивается с необходимостью адекватной интерпретации большого объема полученных данных и их компактной визуализации. Анализируя полученные данные, исследователь замечает тот факт, что оценки изучаемого объекта, полученные по некоторым шкалам, сходны между собой, а если оценить это сходство количественно и подсчитать коэффициент корреляции, то он может оказаться достаточно высоким. Другими словами, возникает вопрос о том, что многие характеристики (т.е. переменные, по которым производилось измерение нашего объекта), вероятно, в некоторой степени дублируют друг друга, а вся полученная информация в целом избыточна. Внимательный исследователь, даже незнакомый с основами ФА, сразу же может сообразить, что за связанными друг с другом (коррелирующими) переменными, по-видимому, стоит влияние некоторой *скрытой, латентной переменной*, с помощью которой можно объяснить наблюдаемое сходство полученных оценок. Очень часто эту гипотетическую латентную переменную называют *фактором.* Приблизительно такая логика заставила Чарлза Спирмена, психолога Оксфордского университета, в ходе анализа результатов тестирования способностей учеников английских школ предположить существование единого, *генерального фактора интеллектуального развития* человека, влияющего на многочисленные показатели разнообразных интеллектуальных тестов. Таким образом, давно известный метод научного познания — обобщение, — приводит нас к возможности и необходимости выделения факторов как переменных более общего, более высокого порядка. Очень часто обобщение позволяет поновому взглянуть на полученные данные, заметить те связи между исходными характеристиками (переменными), которые ранее были не очевидны, а после этого выйти на более высокий уровень понимания сущности измеряемого объекта.

Такого рода обобщение (т.е. сокращение размерности полученных данных) дает возможность использовать очень мощное средство научного анализа — *графическое представление* полученных данных. Понятно, что сокращение размерности результатов многомерного измерения какого-либо объекта до двух-трех позволит исследователю в очень наглядной и компактной форме представить весь объем полученных данных, выйдя за рамки логического анализа массы цифр, собранных в огромную таблицу. Имея в виду важное значение наглядно-образного мышления, трудно переоценить преимущества пространственного (графического) осмысления анализируемых данных. Таким образом, ФА может рассматриваться и как средство компактной визуализации данных.

Выделение в ходе анализа данных общего (для ряда переменных) фактора позволяет решать исследователю еще одну непростую задачу — оценивать некоторую скрытую от непосредственного наблюдения переменную (фактор) опосредованно, косвенно — через ее проявление (влияние) в ряде других, прямо измеряемых переменных. Подобным образом в психодиагностике личности были обнаружены, экстрагированы и измерены многие личностные конструкты, например: классический конструкт Айзенка *импульсивность,* оцениваемый в тесте EPI по ответам испытуемых на ряд вопросов, с разных сторон отражающих этот конструкт. Более того, ФА позволяет измерять не только прямо ненаблюдаемые (скрытые) переменные, но и оценивать определенные качества, которые могут намеренно скрываться и искажаться испытуемым при прямом их тестировании, однако проявляться (т.е. быть измеренными) косвенно через различные связанные с ними качества, оцениваемые прямо.

В ходе научного исследования ФА может выступать в двух ипостасях: как *разведочный (эксплораторный)* и как *проверочный (конфирматорный)* метод анализа данных. В первом случае ФА используется ex post factum, т.е. для анализа уже измеренных в эмпирическом исследовании переменных и, фактически, помогает исследователю их структурировать; на этом этапе совсем не обязательно делать априорные предположения о количестве факторов и их связях с наблюдавшимися переменными. Здесь главное значение ФА — структурировать связи между переменными, помочь сформулировать рабочие гипотезы (пусть иногда и очень умозрительные) о причинах обнаруженных связей. Как правило, такое использование ФА характерно для начальной, ориентировочной стадии работы, когда многое неявное кажется явным, непростое — простым, и наоборот. В отличие от разведочного, конфирматорный ФА используется на более поздних стадиях исследования, когда в рамках какой-либо теории или модели сформулированы четкие гипотезы, связи между переменными и факторами достаточно определены, и исследователь их может прямо указать. Тогда конфирматорный ФА выступает как средство проверки соответствия сформулированной гипотезы полученным эмпирическим данным.

Обобщая вышесказанное, выделим основные цели использования ФА:

1. Понижение размерности числа используемых переменных за счет их объяснения меньшим числом факторов. Обобщение полученных данных.

2. Группировка, структурирование и компактная визуализация полученных данных.

3. Опосредованное, косвенное оценивание изучаемых переменных в случае невозможности или неудобства их прямого измерения.

4. Генерирование новых идей на этапе разведочного анализа. Оценка соответствия эмпирических данных используемой теории на этапе ее подтверждения.

§ 2. Исходные принципы и предположения

Основные общенаучные идеи, лежащие в основе ФА, достаточно просты и могут быть, по мнению П. Благуша (1989), сформулированы так:

а) “сущность вещей заключена в их простых и вместе с тем многообразных проявлениях, которые могут быть объяснены с помощью комбинации нескольких основных факторов”, т.е. за наблюдаемой вариацией достаточно большого количества переменных стоит ограниченное число факторов;

б) “общую сущность наблюдаемых вещей мы постигаем, совершая бесконечные приближения к ней”, т.е. поиск факторов — это длительный процесс познания посредством перехода к факторам все более высокого порядка.

Первым основным формально-математическим принципом, лежащим в основе классической модели ФА[1], является постулат о линейной зависимости между психологическими характеристиками (наблюдаемыми переменными), с помощью которых оценивается какой-либо объект. Количественно степень этой зависимости (связи) может быть оценена с помощью коэффициента корреляции. Второе основное предположение состоит в том, что эти наблюдаемые переменные (предполагается, что их заведомо избыточное количество) могут быть представлены как линейная комбинация некоторых латентных переменных или факторов. Полагается, что ряд этих факторов являются *общими* для нескольких переменных, а другие, *характерные* факторы, специфическим образом связаны только с одной переменной. Поскольку последние ортогональны друг к другу, то, в отличие от общих факторов, они не вносят вклад в корреляцию (ковариацию)[2] между переменными. Таким образом, математическая модель ФА сходна с обычным уравнением множественной регрессии:

---

[1] В данном параграфе мы излагаем наиболее традиционные принципы, лежащие в основе ФА, и принцип линейной зависимости, конечно же, — один из главных. Однако, следует отметить, что в последние годы разрабатываются модели ФА, основанные на более общем предположении — о нелинейной зависимости между наблюдаемыми переменными (Благуш, 1989).

[2] Поскольку ФА работает как с ковариационными, так и с корреляционными матрицами переменных, то мы без особой необходимости не будем подчеркивать различия между ними.

$$V_i = A_{i,1}F_1 + A_{i,2}F_2 + \dots + A_{i,k}F_k + U, \qquad (1)$$

где V_i — значение i-й переменной, которое выражено в виде линейной комбинации k общих факторов, $A_{i,k}$ — регрессионные коэффициенты, показывающие вклад каждого из k факторов в данную переменную; $F_{1\dots k}$ — факторы, общие для всех переменных; U — фактор, характерный только для переменной V_i.

Уравнение (1) выражает весьма простой смысл: каждая переменная может быть представлена в виде суммы вкладов каждого из общих факторов. С другой стороны, аналогичным образом, каждый из k факторов выражается в виде линейной комбинации наблюдаемых переменных:

$$F_j = W_{j,1*}V_1 + W_{j,2*}V_2 + \dots + W_{j,p*}V_p, \qquad (2)$$

где $W_{j,i}$ — нагрузки j-го фактора на i-ю переменную или *факторные нагрузки;* p — количество переменных.

На рис. 1 факторные нагрузки ($w_{1,1} \dots w_{2,6}$) обозначены различными стрелками, показывающими влияние фактора на конкретную переменную. Переменные v1, v2 и v3 преимущественно связаны с фактором F1, и только фактор F2 имеет небольшую нагрузку на первую переменную; для других трех переменных (v4, v5, v6) общим фактором является F2, и лишь на четвертую переменную F1 имеет незначительную нагрузку. Эмпирические оценки наблюдаемых переменных v1 ... v6 представлены в столбцах a, b, c, d, e, f, соответственно. Дугообразная стрелка, соединяющая факторы и коэффициент корреляции над ней, подчеркивают факт *ортогональности* (некоррелированности, линейной независимости) этих факторов, хотя в общем случае (об этом ниже) это предположение критично лишь на этапе выделения первоначальных факторов, а в дальнейшем, на этапе интерпретации факторного решения, при вращении факторной структуры допускается возможность корреляции между факторами. (Это один из многих парадоксов ФА, связанный с *многозначностью* получаемого *факторного решения*, которое не имеет строго однозначного математического обоснования.)

Пользуясь схемой (рис. 1), еще раз обозначим основную задачу ФА: основываясь на эмпирических оценках

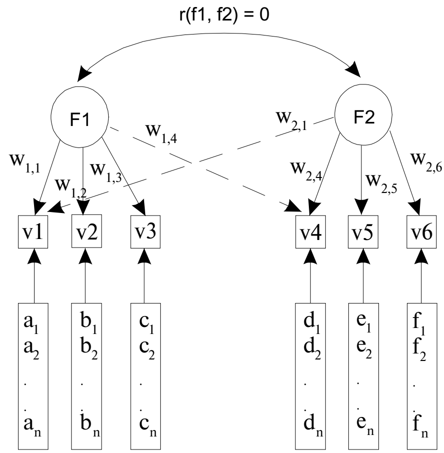


Рис.1. Гипотетическая модель с двумя общими факторами (F1 и F2) и шестью переменными (v1 ... v6)

(a, b, c, d,e, f) исследуемого объекта по каждой из шести переменных-харктеристик (v1 ... v6), исследователь пытается объяснить взаимосвязь наблюдаемых переменных влиянием 2-х общих факторов, в которых находят свое отражение эти переменные.

§ 3. Основные этапы факторного анализа

В ходе исследования с использованием разведочного ФА можно выделить три различных этапа: 1) сбор эмпирических данных и подготовка корреляционной (ковариационной) матрицы; 2) выделение первоначальных (ортогональных) факторов; 3) вращение факторной структуры и содержательная интерпретация результатов ФА. Остановимся на них подробнее.

1. Сбор эмпирических данных в психологическом исследовании разведочного плана всегда опосредован использованием какой-либо измерительной процедуры, в ходе которой испытуемый оценивает измеряемый объект (стимул) по ряду предложенных исследователем характеристик. На этом этапе очень важно, чтобы исследователем был предложен достаточно большой набор характеристик, всесторонне описывающих измеряемый объект. Подбор важных и разнообразных характеристик и одновременно исключение лишних и несущественных — это достаточно трудное дело, требующее от исследователя опыта, знания литературы и, в известной степени, интуиции. Именно продуманный и удачный подбор оцениваемых характеристик определяет в конечном счете успех в выделении существенных и значимых факторов, стоящих за ними — это основное, о чем нельзя забывать на данном этапе. Иначе говоря, из случайного набора характеристик объекта невозможно выделить такие факторы, которые будут закономерно и содержательно определять его оценку испытуемыми. Понятно, что с первого раза, априорно бывает трудно подобрать нужные характеристики. Поэтому еще раз напомним, что разведочное исследование с помощью ФА — это длительный и интеративный процесс, когда результаты предыдущего анализа позволяют оценить допущенные ошибки и скорректировать процедуру последующего исследования.

Второе существенное замечание возникает в связи с постулатом линейности. В случае, когда связь между психологическими характеристиками оказывается существенно нелинейной, базисная размерность искомого факторного пространства возрастает, и это приводит к ложному решению. Преодоление этой трудности может идти двумя путями. Во-первых, можно использовать *коэффициент криволинейной корреляции* (по Пирсону, например), а во-вторых, следует избегать тех психологических переменных, которые имеют между собой явно нелинейные связи.

На данном этапе нельзя не коснуться вопроса о необходимом *уровне измерения,* поскольку он в первую очередь связан с использованием конкретного метода измерения. Вычислительные алгоритмы ФА требуют, чтобы измерения наблюдаемых переменных были проведены не ниже, чем по *шкале интервалов*. Это требование, к сожалению, выполняется далеко не всегда, что, впрочем, связано не столько с неосведомленностью исследователя, сколько с ограниченностью выбора измерительного метода и/или его адекватностью конкретной задаче или даже процедуре исследования. Реалии практики использования ФА в психологии таковы, что в подавляющем большинстве работ применяется один из вариантов метода балльной оценки, который, как известно, дает *шкалу порядка*. Налицо явное ограничение в использовании ФА. При решении данной проблемы следует иметь в виду следующее. Во-первых, стоит уделить максимальное внимание проработке процедурных моментов в использовании метода балльной оценки, чтобы выйти за установление только порядковых отношений и максимально “приблизиться” к шкале интервалов. Во-вторых, следует помнить, что математическая процедура ФА оказывается достаточно устойчивой к разного рода измерительным некорректностям при оценке коэффициентов корреляции между переменными. И наконец, в самой математической статистике существуют различные подходы к решению данной проблемы (Дж. Ким, Ч. Мьюллер,1989), и для более качественной (не строго метрической) трактовки результатов ФА указанное ограничение приобретает не слишком принципиальное значение.

Достаточно важен вопрос о количестве используемых переменных или, более операционально, о том, сколько переменных должно приходиться на один гипотетический фактор. Вслед за Терстоуном многие авторы считают, что в разведочном ФА на один фактор должно приходиться *не менее трех* переменных. Для конфирматорного ФА эта пропорция меньше и, как правило, исследователи ограничиваются двумя переменными. Если исследователя интересует оценка надежности получаемых факторных нагрузок, существуют и более строгие оценки количества необходимых переменных (Дж. Ким, Ч. Мьюллер, 1989).

Формальный итог первого этапа — получение *матрицы смешения* и на ее основе — *корреляционной матрицы*. Матрица смешения — это таблица, куда заносятся результаты измерения наблюдаемых переменных: в столбцах матрицы (по числу переменных) представлены оценки испытуемых (или одного испытуемого) каждой из переменной; строки матрицы — это различные наблюдения каждой переменной. Если задача исследователя — построить факторное пространство для одного испытуемого, то нужно обеспечить множественность таких наблюдений (например, повторить их несколько раз). В том случае, когда строится групповое факторное пространство, достаточно получить по одной оценке от каждого испытуемого. Для последующего расчета по этим данным корреляционной матрицы с достаточно достоверными коэффициентами корреляции следует обеспечить необходимое число наблюдений, т.е. количество строк в матрице смешения. Обычно не следует планировать менее 11—12 наблюдений.

Корреляционная матрица (матрица попарных корреляций между переменными) рассчитывается, как правило, с использованием коэффициента линейной корреляции Пирсона. Часто возникает вопрос о возможности и правомерности использовать другие меры сходства (сопряженнности) между переменными, основанные на ранговой (порядковой) статистике. Понятно, что данный вопрос возникает всегда, когда исследователь работает с номинальными или порядковыми данными. В строгом смысле ответ будет отрицательным. Однако следует принять во внимание два соображения. Во-первых, показано, что при достаточном числе наблюдений коэффициент линейной корреляции Пирсона достаточно устойчив к использованию при расчетах результатов порядковых измерений. Во-вторых, как было отмечено выше, если перед исследователем стоит задача не столько количественного, сколько качественного анализа данных, то такое эвристическое использование ФА считается вполне оправданным.

Еще один тонкий вопрос, связанный с построением матрицы попарных корреляций связан с тем, какую матрицу использовать в ФА — корреляционную или ковариационную? Для начала напомним соответствующие формулы.

Коэффициент ковариации

$$Cov = \frac{1}{n}[\sum(x_i - X)(y_i - Y)] \quad (3)$$

между двумя переменными x и y, а коэффициент корреляции:

$$r_{xy} = Cov/s_x s_y \quad , \quad (4)$$

где n — количество наблюдений, x_i и y_i — значения переменных x и y; X и Y — средние арифметические значения переменных x и y по ряду наблюдений; σ_x и σ_y — средние квадратические отклонения переменных x и y по ряду наблюдений.

Таким образом очевидно, что коэффициент корреляции — это тот же коэффициент ковариации, только нормированный по среднему квадратическому отклонению или, как еще говорят, выраженный в единицах среднего квадратического отклонения переменных. Из этого следуют и "рецепты" по применению в ФА корреляционной или ковариационной матриц:

1) если все переменные выражены в одних и тех же единицах измерения, то нет большого различия, какую из матриц факторизовать;

2) если метрики переменных заметно отличаются (единицы измерения значительно неоднородны и дисперсии переменных заметно отличаются), то целесообразно использовать анализ корреляционной матрицы;

3) ковариационные матрицы предпочтительнее, когда необходимо провести сравнение результатов ФА (факторных структур) в двух различных выборках, полученных в одном и том же исследовании, например, когда требуется оценить повторяемость какого-либо интересного результата.

2. Следующий важнейший этап ФА — собственно факторизация матрицы корреляций (ковариаций) или выделение первоначальных (ортогональных) факторов. В настоящее время — это полностью компьютеризованная процедура, которую можно найти во всех современных статистических программах. Одним из первых, кто предложил формально-математическое решение проблемы возможности факторизации корреляционной матрицы, был Л. Терстоун. В матричной форме его известное уравнение выглядит следующим образом (подробнее см.: Я. Окунь, 1974, с. 43—49):

$$\|R\| = \|F\| \times \|F'\| , \tag{5}$$

где $\|R\|$ — редуцированная корреляционная матрица;

$\|F\|$ — редуцированная матрица факторных нагрузок;

$\|F'\|$ — транспонированная матрица факторных нагрузок.

Поясним, что *редуцированная корреляционная матрица* — это матрица попарных корреляций наблюдаемых переменных, где на главной диагонали лежат не единицы (как в полной матрице корреляций), а значения, соответствующие влиянию только *общих* для этих переменных факторов и называемые *общностями*. Аналогичным образом, редуцированная матрица факторных нагрузок или факторная матрица (формальная цель ФА) представляет собой факторные нагрузки только общих факторов.

Основная проблема, стоящая при решении уравнения (3), заключается в том, что значения общностей в редуцированной корреляционной матрице неизвестны, а для начала вычислений их необходимо иметь. На первый взгляд неразрешимая проблема решается таким образом: до начала вычислений задаются некоторые приблизительные значения общностей (например, максимальный коэффициент корреляции по столбцу), а затем на последующих стадиях вычислений, когда уже имеются предварительные величины вычисленных факторных нагрузок, они уточняются. Таким образом, очевидно, что вычислительные алгоритмы ФА представляют собой последовательность итеративных[1] вычислений, где результаты каждого последующего шага определяются результатами предыдущих. С известной долей упрощения можно считать, что различ-

---

[1] Итерация — это математический термин, означающий результат применения какой-либо математической операции, получающийся в серии аналогичных операций.

ные алгоритмы факторизации корреляционной матрицы в основном и отличаются тем, как конкретно решается данная проблема.

Для людей, неискушенных в проблемах математической статистики, но решающих с помощью ФА свою задачу, более важен основной смысл процедуры факторизации, заключающийся в переходах от матрицы смешения к корреляционной матрице и далее к матрице факторных нагрузок и построению факторных диаграмм (рис. 2).

Матрица смешения

N	v1	v2	v3	v4	v5	v6
1	a_1	b_1	c_1	d_1	e_1	f_1
2	a_2	b_2	c_2	d_2	e_2	f_2
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
n	a_n	b_n	c_n	d_n	e_n	f_n

⇒ Корреляционная матрица

$r_{1,1}$	$r_{1,2}$	$r_{1,3}$	$r_{1,4}$	$r_{1,5}$	$r_{1,6}$
$r_{2,1}$	$r_{2,2}$	$r_{2,3}$	$r_{2,4}$	$r_{2,5}$	$r_{2,6}$
$r_{3,1}$	$r_{3,2}$	$r_{3,3}$	$r_{3,4}$	$r_{3,5}$	$r_{3,6}$
$r_{4,1}$	$r_{4,2}$	$r_{4,3}$	$r_{4,4}$	$r_{4,5}$	$r_{4,6}$
$r_{5,1}$	$r_{5,2}$	$r_{5,3}$	$r_{5,4}$	$r_{5,5}$	$r_{5,6}$
$r_{6,1}$	$r_{6,2}$	$r_{6,3}$	$r_{6,4}$	$r_{6,5}$	$r_{6,6}$

⇒ Факторная матрица

v_i	F1	F2
v1	$w_{1,1}$	$w_{1,2}$
v2	$w_{2,1}$	$w_{2,2}$
v3	$w_{3,1}$	$w_{3,2}$
v4	$w_{4,1}$	$w_{4,2}$
v5	$w_{5,1}$	$w_{5,2}$
v6	$w_{6,1}$	$w_{6,2}$

⇓

Факторная даграмма

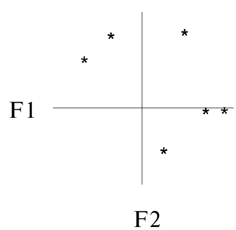


Рис. 2. Основные этапы трансформации данных в ходе факторного анализа.

Пользуясь данным рисунком, еще раз подчеркнем важную особенность ФА — это способ понижения размерности, сжатия объема данных. Обратите внимание, что исходная матрица смешения достаточно велика например, при условии 20-ти наблюдений каждой переменной она содержит 20×6=120 измерений. Конечный результат анализа — это всего лишь 2×6=12 чисел или построенная по матрице факторных нагрузок компактная факторная диаграмма. Таким образом, при адекватном использовании ФА как метода многомерного измерения мы можем получить 10-кратную компрессию исходной информации и наглядность результатов ее анализа.

Напомним, что главная цель выделения первичных факторов в разведочном ФА состоит в определении *минимального* числа общих факторов, которые удовлетворительно воспроизводят (объясняют) корреляции между наблюдаемыми переменными. Основная стратегия при выделении факторов незначительно отличается в разных методах. Она заключается в оценке гипотезы о минимальном числе общих факторов, которые оптимально воспроизводят имеющиеся корреляции. Если нет каких-либо предположений о числе факторов (в ряде программ оно может быть задано прямо), то начинают с однофакторной модели. Эта гипотеза о достаточности одного фактора оценивается с помощью используемого критерия оптимальности соответствия данной однофакторной модели исходной корреляционной матрице. Если расхождение статистически значимо, то на следующем шаге оценивается модель с двумя факторами и т. д. Такой процесс *подгонки* модели под данные осуществляется до тех пор, пока с точки зрения используемого критерия соответствия расхождение не станет минимальным и будет оцениваться как случайное. В современных компьютерных статистических программах используются различные методы факторизации корреляционной матрицы. Нам представляется, что, хотя для исследователя данная проблема не представляет прямого интереса, тем не менее она важна, поскольку от выбора метода факторизации в определенной мере зависят результаты расчета факторных нагрузок. В силу специфики нашего изложения основ ФА мы ограничимся лишь перечислением этих методов, снабдив его очень краткими комментариями и отошлем читателя для более глубокого знакомства к специальной литературе, требующей некоторых познаний в математике (Дж. Ким, Ч. Мьюллер, 1989):

М*етод главных факторов* (или главных осей) — наиболее старый и часто используемый в различных предметных областях.

Метод наименьших квадратов сводится к минимизации остаточной корреляции после выделения определенного числа факторов и к оценке качества соответствия вычисленных и наблюдаемых коэффициентов корреляции по критерию минимума суммы квадратов отклонений.

Метод максимального правдоподобия: специфика данного метода состоит в том, что в случае большой выборки (большого количества наблюдений каждой переменной) он позволяет получить статистический критерий значимости полученного факторного решения.

Альфа-факторный анализ был разработан специально для анализа психологических данных, и поэтому его выводы носят в основном психометрический, а не статистический характер. В альфа-факторном анализе минимальное количество общих факторов оценивается по величинам *собственных значений* факторов и *коэффициентов обобщенности* α, которые должны быть больше 1 и 0, соответственно.

Факторизация образов (или анализ образов). В отличие от классического ФА в анализе образов предполагается, что общность каждой переменной определяется не как функция гипотетических факторов, а как линейная регрессия всех остальных переменных.

В табл. 1 представлены сравнительные результаты факторизации корреляционной матрицы (Дж. Ким, Ч. Мьюллер, 1989, с. 10), с использованием 4-х различных методов. Видно, что полученные результаты могут различаться, даже если не обращать внимание на знаки факторных нагрузок (об этом чуть ниже).

После компьютерного расчета матрицы факторных нагрузок наступает наиболее сложный, ответственный и творческий этап использования ФА — определение минимального числа факторов, адекватно воспроизводящих наблюдаемые корреляции, и содержательная интерпретация результатов ФА. Напомним, что максимальное количество факторов равно числу переменных. Кроме со-

Таблица 1

Использование различных методов факторизации для получения двухфакторного решения

Переменная	Метод главных факторов		Метод макс. правдоподобия		Альфа факторный анализ		Анализ образов	
	F1	F2	F1	F2	F1	F2	F1	F2
v1	0.73	— 0.32	0.75	— 0.30	0.70	0.44	0.58	0.13
v2	0.64	— 0.28	0.70	— 0.27	0.59	0.38	0.54	0.14
v3	0.55	— 0.24	0.60	— 0.18	0.50	0.33	0.48	0.13
v4	0.51	0.47	0.43	0.36	0.59	— 0.38	0.37	— 0.27
v5	0.44	0.41	0.51	0.61	0.50	— 0.33	0.39	— 0.26
v6	0.37	0.34	0.53	0.25	0.42	— 0.27	0.29	— 0.24

держательных критериев решения вопроса о минимальном числе факторов существуют формально-статистические показатели *достаточности* числа выделенных факторов для объяснения корреляционной матрицы. Остановимся на двух основных показателях. После расчета факторных нагрузок для каждой переменной практически любая компьютерная программа распечатывает на экране следующую табл. 2[1].

Первый важный показатель в этой таблице (второй столбец) — это величина *собственного значения* каждого фактора; факторы расположены в таблице по убыванию этой величины. Не очень строго говоря, этот показатель характеризует *вес*, значимость каждого фактора в найденном факторном решении[2]. Из таблицы 2 видно, что от 1-го фактора к 10-му (всего было 10 переменных) величи-

[1] Результаты ФА взяты из работы студентов 2-го курса д/о ф-та психологии Московского государственного университета им. М.В. Ломоносова, 1995/96 уч. год.

[2] Более точно, собственное значение каждого фактора — это его вклад в дисперсию переменных, объясняемую влиянием общих факторов. Расчет собственных значений корреляционной матрицы — один из основных вычислительных алгоритмов ФА.

Таблица 2

Статистические показатели для определения минимального количества факторов

Фактор	Собственное значение	% объясняемой дисперсии	Сумм. % объясняемой дисперсии
1	5,14	51.4	51.4
2	1.72	17.2	68.6
3	1.03	10.3	78.9
4	0.76	7.7	86.6
5	0.38	3.9	90.5
6	0.33	3.3	93.7
7	0.28	2.8	96.6
8	0.21	2.1	98.7
9	0.08	0.8	99.5
10	0.05	0.5	100

на собственного значения убывает более, чем в 100 раз. Естественно возникает вопрос о том, какая величина данного показателя свидетельствует о значимом, существенном вкладе соответствующего фактора, и каков критерий для отсечения незначимых, несущественных факторов. Достаточно часто в качестве такого критического значения используют величину собственного значения, равную 1.0. Таким образом, с определенной степенью уверенности предполагают, что те факторы, у которых этот показатель меньше 1.0, не вносят значительного вклада в объяснение корреляционной матрицы. Кроме анализа табличных величин бывает полезно визуально оценить динамику величины собственного значения по графику. Как правило, в большинстве статистических программ такая возможность пользователю предоставляется (см. рис. 3). Как предлагал в свое время Р. Кеттел (1965), выделение факторов заканчивается, когда после резкого падения величины собственного значения изменяются незначительно, и график фактически превращается в горизонтальную прямую линию. Несмотря на видимую простоту и ясность такого рецепта, следует отметить, что когда на графике имеется более чем один излом, то выделение горизонтального участка становится неоднозначным.

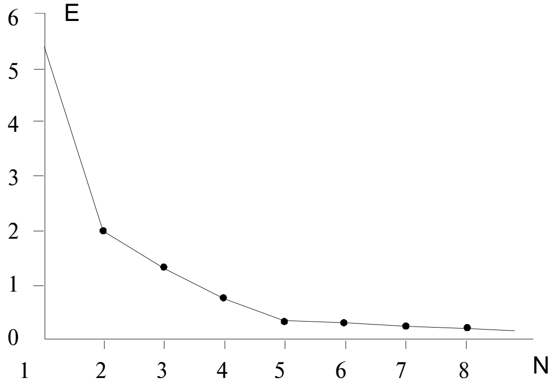


Рис. 3. Изменение величины собственного значения факторов

Другой не менее важный расчетный показатель значимости каждого фактора — *процент объясняемой дисперсии переменных*, содержащейся в корреляционной матрице (третий столбец в табл. 2). Естественно, что все 100% дисперсии будут объясняться только всеми десятью факторами. Однако не стоит забывать, что при любых измерениях (а особенно в разведочных, пилотажных исследованиях) имеют место разного рода случайные ошибки, и поэтому их вклад в общую дисперсию тоже может оказаться весьма значительным. Предполагается, что ряд выделенных факторов отражает влияние случайных процессов, никак не связанных с оценкой наблюдаемых переменных. Таким образом, формально задача заключается в том, чтобы, с *одной стороны, выбрать некоторое минимальное количество факторов, которые бы, с другой стороны, объясняли достаточно большой процент всей дисперсии переменных.* Ясно, что эти два требования в принципе взаимно противоречивы, и, следовательно, исследователь стоит перед выбором некоторой критической величины процента объясняемой дисперсии. К сожалению, никаких строго формальных рецептов по этому поводу не существует, но принято считать, что при хорошем факторном решении выбирают столько факторов, чтобы они в сумме (последний столбец таблицы) объясняли не менее 70—75%. В хорошо спланированных исследованиях с установленной факторной структурой этот *суммарный процент* может достигать 85—90 %.

Подводя итог, укажем, что в данном случае оба статистических критерия вполне однозначно свидетельствуют о достаточности не более 3-х факторов, что и отмечено пунктирной горизонтальной линией. Тем не менее, стоит подчеркнуть, что главным критерием для выделения минимального количества будет содержательная интерпретация выделенных факторов, а к использованию формально-статистических критериев следует относиться с осторожностью.

3. Вращение факторной структуры и содержательная интерпретация результатов ФА. Одним из основных кажущихся парадоксов ФА как метода, обеспеченного весьма солидным и современным математическим аппаратом, является *неоднозначность расчета факторных нагрузок* по исходной корреляционной матрице. Фактически это означает следующее: любой алгоритм факторизации корреляционной матрицы дает какой-то один вариант расчета факторных нагрузок из целого множества эквивалентных. Это означает, что расчет факторных нагрузок выполняется с точностью до *любого* линейного преобразования в правой части уравнения (2), что эквивалентно возможности произвольного поворота факторных осей вокруг векторов-переменных. Поясним сказанное, используя геометрическую интерпретацию результатов ФА. На рис. 4 представлены три переменные (v1, v2 и v3) в пространстве двух ортогональных факторов (F1 и F2). Переменные изображены в виде векторов, а факторные нагрузки переменных на факторы геометрически представляют собой проекции данного вектора (переменной) на соответствующую координатную ось (фактор). Если мы осуществим произвольный поворот осей координат на какой-то угол, например, на 45 градусов вправо (новые оси — штрих-пунктирные линии на рис. 4), то расположение переменных в новой системе координат

(*F1′— F2′*) с математической точки зрения полностью эквивалентно исходному. Изменились лишь величины факторных нагрузок (сравните проекции переменной v1 на оси F1 и *F1′*, соответственно, до и после поворота). Таким образом, исходное факторное решение справедливо с точностью до любого угла поворота ортогональных факторных осей вокруг пучка векторов, образованного переменными v1, v2 и v3.

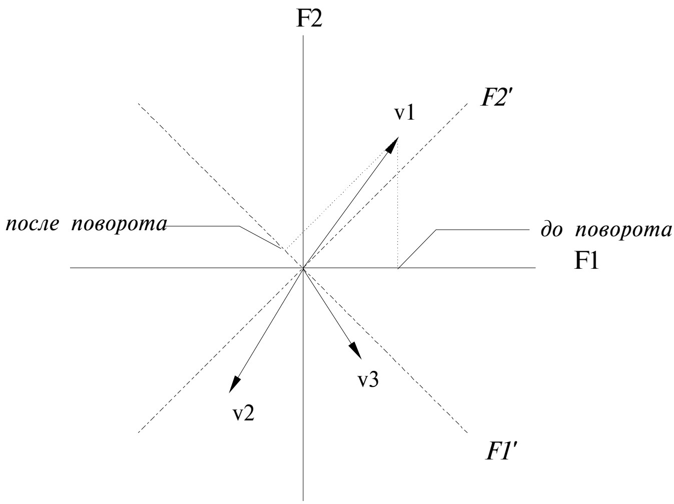


Рис.4. Факторное пространство 3-х переменных (v1, v2 и v3) в пространстве 2-х факторов:
сплошные линии (F1, F2) — до поворота; пунктирные линии (*F1′*, *F2′*) — после поворота

Естественно, возникает вопрос об оптимальном расположении переменных в пространстве факторных осей. Как было отмечено выше, эта проблема в принципе не имеет строгого математического решения. Она относится уже к содержательной интерпретации расположения переменных в факторном пространстве. Фактически суть проблемы состоит в следующем: какой набор факторных нагрузок (или какая геометрическая модель результатов ФА) будет более подходящим для интерпретации исследователем. Поскольку при повороте осей координат факторные нагрузки по одному фактору могут расти, а по другому — уменьшаться, то, соответственно, будет расти или уменьшаться вклад этих факторов в разные переменные. Из этого следует, что нужно искать такой вариант расположения переменных в факторном пространстве, который наилучшим образом соответствует ожиданиям исследователя, его предположениям о взаимосвязи и взаимозависимости исследуемых переменных. Как правило, при использовании ФА полагают, что существует одно оптимальное положение осей координат, соответствующее существенным для данного исследования и хорошо интерпретируемым факторам.

Описанный выше процесс поиска оптимальной факторной структуры получил название процедуры *вращения факторов.* По образному выражению Л. Терстоуна, цель исследователя заключается в поиске *“простой структуры”* или попытке объяснить большее число переменных меньшим числом факторов. С формальной точки зрения при поиске простой структуры следует иметь в виду следующее: целесообразно стремиться к получению для каждой переменной максимального числа *больших* факторных нагрузок *по одним факторам* и одновременно наибольшего количества *минимальных* факторных нагрузок *по другим факторам.* Следуя этому правилу, мы стремимся сделать так, чтобы одну группу переменных можно было в большей степени объяснить влиянием одних факторов, а другую — других. Таким образом, “простота” хорошего факторного решения заключается в том, что каждая переменная имеет наиболее простое факторное объяснение, т.е. характеризуется преобладающим влиянием некоторого одного фактора, и в меньшей степени связана с другими факторами. И наоборот: один фактор должен быть специфическим образом связан с одной группой переменных и не связан с другими переменными. В предельном случае самая простая структура получается тогда, когда все переменные располагаются на соответствующих факторных осях, т. е. имеют ненулевые факторные нагрузки только по одному фактору, а по остальным — нулевые. Возвращаясь к рис. 4, укажем на результат вращения: после поворота факторных осей на 45 градусов вправо, нагрузка переменной v1 по первому фактору резко уменьшилась и одновременно немного возросла по второму. Кроме того заметно уменьшились факторные нагрузки переменных v2 и v3 по второму фактору. Таким образом “простота” новой факторной структуры выразилась в доминирующем влиянии первого фактора на переменную v1, а второго фактора — на две других переменных.

Видимая простота решения проблемы вращения системы координат в двухмерном случае (это можно сделать и вручную) становится неочевидной при 3-х и более факторах. Пересчет системы координат вручную и построение факторных диаграмм становятся очень сложными. Исходя из общего принципа построения простой структуры, изложенного выше, во многих компьютерных программах предлагаются несколько способов решения проблемы оптимальности вращения системы координат. Кратко остановимся на основных способах вращения. Выделяют два класса методов вращения — методы *ортогонального* вращения, когда при повороте осей координат угол между факторами остается прямым (и, следовательно, остается верным предположение о некоррелированности факторов), и более общие методы *косоугольного* вращения, когда первоначальное ограничение о некоррелированности факторов снимается.

Методы ортогонального вращения: варимакс, квартимакс, эквимакс и биквартимакс. *Варимакс* — наиболее часто используемый на практике метод, цель которого — минимизировать количество переменных, имеющих высокие нагрузки на данный фактор, что способствует упрощению описания фактора за счет группировки вокруг него только тех переменных, которые с ним связаны в большей степени, чем остальные.

Квартимакс в определенном смысле противоположен варимаксу, поскольку минимизирует количество факторов, необходимых для объяснения данной переменной; поэтому он усиливает интерпретабельность переменных. Квартимакс-вращение приводит к выделению одного из общих факторов с достаточно высокими нагрузками на большинство переменных. *Эквимакс и биквартимакс* — это два схожих метода, представляющих собой своеобразную комбинацию варимакса, упрощающего описание факторов, и квартимакса, упрощающего описание переменных. Отметим, что выбор более подходящего метода вращения конечно же требует известного опыта использования ФА, однако специальные исследования Х. Кайзера (1958) при прочих равных условиях свидетельствуют в пользу преимущественного применения варимакса.

Методы косоугольного вращения также позволяют упростить описание факторного решения за счет введения предположения о коррелированности факторов и, следовательно, о возможности существования *факторов более высокого порядка*, объясняющих наблюдаемую корреляцию. Основное преимущество косоугольного вращения состоит в возможности проверки ортогональности получаемых факторов: если в результате вращения получаются действительно ортогональные факторы, то можно быть уверенным в том, что ортогональность им действительно свойственна, а не является следствием использования метода ортогонального вращения. В статистических программах наибольшую популярность получил метод *облимин,* по своей сути эквивалентный методу эквимакс при ортогональном вращении. В расчетах с помощью облимина используется специальный параметр (называемый в разных программах α или δ), задающий степень косоугольности факторов при вращении. Большие отрицательные значения этого параметра соответствуют наиболее косоугольным решениям, а меньшие отрицательные значения — наиболее ортогональному решению. Подробнее об использовании метода облимин можно прочитать в книге Г. Хармана (1972) и руководствах к соответствующим статистическим пакетам.

Стоит особо отметить, что перед выполнением процедуры вращения следует указать количество факторов, в пространстве которых и производится вращение. Поэтому вопрос о минимальном количестве факторов следует решить (в первом приближении!) до того. После осуществления вращения и анализа факторных диаграмм можно еще раз вернуться к проблеме минимального количества факторов, чтобы затем еще раз выполнить вращение с другим количеством (меньшим или большим) факторов. Например, на основании использования ряда статистических критериев, описанных выше, мы начинаем проводить вращение с учетом наличия 4-х факторов, но в ходе анализа факторных диаграмм убеждаемся в избыточности третьего и четвертого факторов и окончательное вращение выполняем в 2-х факторном пространстве. Таким образом, вращение и анализ факторных диаграмм следует проводить несколько раз с учетом разного количества факторов, начиная, как правило, с избыточного количества.

Вместе с тем не следует думать, что получение простой геометрической модели факторного решения является основным критерием "хорошести" результатов ФА и, следовательно, единственности и оптимальности расположения исследуемых переменных в системе факторных координат. Безусловно, решение вопроса о минимальном количестве факторов и сравнительная оценка различных вариантов вращения должны основываться на серьезном содержательном анализе полученных результатов. Укажем на основные моменты в ходе содержательного анализа:

1. По возможности следует четко сформулировать ожидаемые результаты и после этого задать себе следующие вопросы: а) согласуются ли ваши данные с вашими ожиданиями и результатами ранее выполненных исследований? б) что общего и какие есть различия между вашим и другими подобными исследованиями?

2. Стоит вспомнить, использовался ли ФА в сходных исследованиях и какие факторы выделялись в таких работах.

3. И наконец, при интерпретации факторов и объяснении их влияния на исследуемые переменные, следует подумать о согласованности найденного вами факторного решения с теоретическими основаниями данной предметной области психологии.

§ 4. Дополнительные статистические показатели для оценки результатов факторного анализа

В начале предыдущего параграфа мы отмечали, что вычислительные алгоритмы ФА основываются на ряде математических допущений о характере эмпирических данных, подвергаемых ФА. Остановимся на ряде статистических показателей, которые помогают исследователю оценить степень соответствия данных этим допущениям.

Как правило, в любой программе по ФА предусмотрен расчет показателей описательной статистики по матрице смешения. Например в статистических системах "Stadia" и SPSS для каждой переменной вычисляются общее количество наблюдений, среднее арифметическое значение и среднее квадратичное отклонение (см. табл. 3). Эти достаточно простые показатели позволяют быстро сравнить между собой все анализируемые переменные, и уже на уровне анализа исходных данных попытаться найти возможные ошибки, связанные либо с проведенными измерениями, либо с вводом данных в компьютер. Например, если при сборе данных использовалась 7-балльная шкала, то наверное вас насторожит среднее значение по какой-то переменной, равное 0.87, или резко отличающаяся от других величина среднего квадратичного отклонения.

Коэффициент сферичности Бартлета используется для оценки "хорошести" корреляционной матрицы. Если этот коэффициент достаточно большой, а соответствующий ему уровень значимости мал (например, меньше 0.05 или 0.01), то это свидетельствует о надежности вычисления корреляционной матрицы. При высоком уровне значимости исследователю стоит задуматься об адекватности использования ФА с полученными данными.

Кроме того, для оценки надежности вычислений элементов корреляционной матрицы и возможности ее описания с помощью ФА во многих статистических программах применяется так называемая *мера адекватности вы-*

борки Кайзера—Мейера—Олкина(КМО)[1]. По мнению Г. Кайзера (1974), значения КМО около 0.9 оцениваются как "изумительные", 0.8 — "достойные похвалы", 0.7 — "средние", 0.6 — "посредственные", 0.5 — "плохие", а ниже 0.5 — "неприемлемые". Для оценки надежности вклада в корреляционную матрицу *каждой переменной в отдельности* также используют меру выборочной адекватности (например, коэффициент MSA_i в системе SPSS). Вышеприведенные характеристики Г. Кайзера вполне справедливы и для оценки этих величин тоже. Исследование надежности каждой переменной позволяет исключить из расчетов одну или несколько переменных, и тем самым повысить результативность ФА.

Таблица 3

Данные описательной статистики для 9 переменных

Переменная	Среднее Арифм.	Ср. кв. откл.
F_MMPI	8,89	4,69
K_MMPI	15,36	3,54
L_MMPI	4,30	2,59
MMPI_0	31,30	9,41
MMPI_1	15,89	4,07
MMPI_2	27,38	5,08
MMPI_3	24,88	5,47
MMPI_4	24,78	5,03
MMPI_5	33,97	2,99
MMPI_6	12,41	3,95
MMPI_7	31,22	5,30
MMPI_8	31,83	6,90
MMPI_9	19,19	4,30
Количество наблюдений = 36		

[1] Имеется в виду адекватность факторной модели данному набору переменных, описываемому корреляционной матрицей.

Работая с различными данными, Г. Кайзер установил, что величина данного коэффициента адекватности повышается при: а) увеличении количества переменных, б) возрастании числа наблюдений каждой переменной, в) уменьшении числа общих факторов и г) увеличении абсолютных значений коэффициентов корреляций. По сути дела данный автор выделил те условия, при которых повышается адекватность данных, а следовательно, и информативность ФА.

§ 5. Несколько замечаний по поводу конфирматорного ФА

Как было отмечено выше, конфирматорный ФА используется для проверки и подтверждения теоретической модели факторного типа эмпирическими данными. Предполагается, что у исследователя есть достаточно строго сформулированная модель изучаемой им реальности (например, какие психологические факторы в межкультурном исследовании мотивации достижения у школьников являются общими для всех культур, а какие специфическим образом влияют на мотивационные переменные только в одной стране).

При использовании конфирматорного ФА исследователь (в рамках своей модели) четко формулирует гипотезу о числе общих и специфических факторов. Естественно, эта гипотеза должна основываться на серьезном анализе природы исследуемых переменных и лежащих в их основе факторов. Более того, проверяя свою модель на реальных данных, исследователь может делать и количественные предположения о величине корреляции между переменными, величинах факторных нагрузок для ряда исследуемых переменных и зависимости между факторами (ортогональными или косоугольными). Имея данные эмпирических измерений, с одной стороны, и набор разнообразных теоретических гипотез — с другой, психолог с помощью ФА фактически занимается проверкой сформулированных им гипотез о свойствах изучаемой (моделируемой) реальности.

Подробное изложение исследовательских стратегий с помощью конфирматорного ФА не входит в задачу настоящего учебного пособия, поскольку представляет собой особую и достаточно специфическую задачу. Тем не менее, укажем, что в настоящее время существуют достаточно удобные компьютерные программы, где реализованы современные подходы к анализу моделей с латентными переменными, частным случаем которых и является ФА. В качестве примера мы можем привести достаточно известный статистический пакет *Lisrel 8*, который дает возможность обрабатывать данные *методом моделирования с помощью линейных структурных уравнений*. Для подробного знакомства с принципами конфирматорного ФА могут быть рекомендовано (Благуш, 1989), а также прекрасное описание статистического пакета Lisrel 8[1].

Методические рекомендации по выполнению учебного задания по теме «Факторный анализ»

Основная трудность при выполнении настоящего учебного задания — это, как ни странно, выбрать подходящий предмет для исследования, т.е. определить тот набор переменных, которые необходимо или интересно изучить с помощью ФА. При решении этой проблемы можно пойти двумя путями: либо взять заведомо подходящую задачу, которая ранее уже решалась с помощью ФА, либо придумать ее самому (последнее, естественно труднее, но интереснее). В принципе и то, и другое подходит для выполнения учебного задания. Достаточно стандартный вариант выполнения работы — это провести ФА какого-либо из-

[1] Для знакомства с использованием данного метода в психологии мы советуем прочесть статью Е.Л. Григоренко. Применение статистических методов моделирования с помощью линейных структурных уравнений в психологии: За и Против//Вопросы психологии, 1994. № 4.

вестного опросника, в котором уже содержатся шкалы (факторы) и отражающие их вопросы (переменные). Еще лучше взять какой-либо новый (например, только что переведенный), но еще не стандартизированный опросник и провести исследование с ним. В этом случае будет интересно подумать над интерпретацией результатов ФА, и хотя бы немного побыть в роли разработчика новых психодиагностических методик. Неплохой вариант, если вы найдете в литературе данные, которые можно обработать ФА, и подумаете над их интерпретацией в контексте обсуждаемых автором работы проблем.

Для ориентировки студентов в том, что же можно сделать, мы приводим ниже список названий работ по ФА, которые были выполнены студентами 2-го курса факультета психологии Московского государственного университета им. М.В. Ломоносова в 1995—1996 гг.[1]:

1. Оценка эмоционального состояния при прослушивании музыки разных жанров.

2. Личностные особенности деятелей тайных обществ первой трети XIX века.

3. Факторное пространство русских писателей XIX века.

4. Исследование факторов, определяющих положение человека в семье.

5. Изучение влияния различных типов стрессогенных ситуаций на интенсивность эмоционального переживания: определение специфики ситуаций для мужской и женской выборок.

6. Выделение скрытых факторов, обуславливающих привлекательность печатной рекламы.

7. Выявление факторов, оказывающих наибольшее влияние на выбор того или иного политического лидера при голосовании.

8. Факторы, способствующие заинтересованности человека той или иной пластинкой по виду ее конверта.

9. Характеристика человека, с которым мы хотим дружить.

---

[1] Названия тем студенческих работ приводятся без изменений и редакции.

10. Факторизация шкал опросника “16 PF”.

11. Оценка изучаемых предметов студентами 2-го курса.

12. Выявление факторной структуры шкал акцентуации характера по Леонгарду (тест Шмишека). Сравнение результатов факторизации на 2-х выборках испытуемых.

13. Исследование факторов, влияющих на выбор страны для зарубежной поездки.

14. Факторы, определяющие оценку идеального мужчины и идеальной женщины.

15. Исследование факторов, определяющих специфику национального характера.

После того, как уже выбрана адекватная исследовательская или практическая задача (предмет исследования), которая будет решаться с помощью ФА, и в основном определен набор оцениваемых переменных, стоит еще раз подумать о правильности их выбора. В первую очередь следует обратить внимание на то, чтобы переменные не повторяли друг друга, а разнообразно и всесторонне описывали предмет вашего исследования. В разведочном исследовании тщательный и вдумчивый подбор наблюдаемых переменных может обеспечить *полноту описания* изучаемой реальности. От этого и будет зависеть, сумеете ли вы выделить действительно важные факторы, влияющие на восприятие, оценку, понимание или действия человека в определенной ситуации, описываемой используемыми переменными. Например, если вы решили исследовать психологические факторы, которые определяют восприятие избирателями лидеров политических партий, то не следует ограничиваться оценкой только их личностных особенностей, безусловно стоит включить также и описательные характеристики их внешних данных, политических ориентаций и многое другое. Не следует забывать о том, что исследуемые вами факторы есть не более чем “экстракт” наблюдаемых переменных, и, следовательно, они не могут появиться из ничего.

Однако не стоит и чрезмерно увеличивать число используемых переменных путем включения нескольких однотипных. Если несколько выбранных вами переменных похожи друг на друга, то очевидно, что это приведет к появлению очень высоких коэффициентов корреляции между этими переменными и, таким образом, к *избыточности и односторонности* описания предмета вашего исследования.

В том случае, когда вы затрудняетесь или сомневаетесь в выборе необходимых переменных, полезно создать их заведомо избыточный список, а затем, воспользовавшись правилом “со стороны виднее”, попросить своих коллег поучаствовать в оценке этого списка в качестве экспертов.

Следующий важный этап в проведении исследования — *сбор данных.*

На этом этапе, как правило, сталкиваются с двумя вопросами: по какой группе испытуемых собирать данные и каким методом это делать? На первый вопрос ответить достаточно просто: чтобы получить статистически достоверные оценки коэффициентов корреляции, нужно по каждой переменной собрать *не менее 12—15* наблюдений. Если задача состоит в построении факторного пространства для одного испытуемого, то нужно решить, каким образом лучше получить от него такое количество повторных данных.

При решении второго вопроса мы советуем обратиться к соответствующей главе настоящего пособия, посвященной методу балльной оценки. Какой процедурой сбора данных лучше воспользоваться, зависит от задачи вашего исследования, от условий, в которых проводится тестирование, от возраста и уровня образования испытуемых и т. д. При выборе конкретного варианта методики не стоит забывать и о простоте последующей обработки исходных данных, и об удобстве их считывания с бланка и ввода в компьютер.

Ввод данных и их обработка.

Остановимся кратко на некоторых важных этапах работы со статистической программой, с помощью которой собственно и реализуется процедура ФА. Для этой цели мы рекомендуем использовать либо русскоязычную статистическую систему “Stadia” или англоязычную систему обработки и анализа данных SPSS. Эти две программы достаточно широко используются, соответственно, российскими и зарубежными психологами и ориентированы на пользователя-гуманитария. Для облегчения использования этих двух программ, мы остановимся на основных моментах работы с каждой из них.

Работа в системе "Stadia". После вызова программы (stadia.exe) вы сразу же попадаете в редактор данных и, следовательно, можете начинать ввод данных в электронную таблицу. Закончив ввод данных, не забудьте их *сохранить на жестком диске*— **F4** ; практика показывает, что несоблюдение этого правила для неопытного пользователя часто заканчивается повторным вводом данных. Кроме того, обязательно проверьте правильность ввода данных (лучше эту малоприятную процедуру выполнять вдвоем: один читает — другой проверяет). В том случае, если данные уже набраны в каком-либо текстовом редакторе, вы можете загрузить их в окно редактора с дискеты, для чего используйте функцию "Чтение" — **F3.**

Войдя в меню статистических методов (**F9**), выберите в разделе "Многомерные методы" опцию "Факторный анализ". Первый запрос программы касается типа введенных данных — что это: матрица смешения (переменные объекты) или корреляционная матрица; как правило, вы начинаете работать с матрицей смешения. После рассчета корреляционной матрицы появляется вопрос: "Записать ли рассчитанные корреляции в матрицу данных?"; чаще всего в этом нет особой необходимости. Далее на экране распечатывается таблица с показателями описательной статистики и матрица корреляций. Эта уже та информация, которую стоит записать в файл результатов — **F2**; в качестве имени файла (без расширения!) целесообразно ввести первые 6—8 букв своей фамилии латинскими буквами. Если выводимая на экран информация не уместилась на одной экранной странице, нажмите клавишу "Enter". После этого на экране распечатается таблица с величинами собственных значений и процентом объясняемой дисперсии факторов (не забудьте сохранить и ее!) и появляется вопрос: "Выдать собственные векторы и новые координаты объектов?”; поскольку анализ собственных векторов используется редко, ответьте — “нет”. А вот график собственных значений посмотреть весьма полезно, поэтому на следующий вопрос программы ответьте “да” и посмотрите его *на экране*. Затем производится расчет первичных факторных нагрузок и соответствующая матрица распечатывается на экране. Можно ее сохранить в файле и посмотреть факторные диаграммы, а можно ответить “нет” (чаще всего так и поступают) и, нажав “Enter”, сразу перейти к вращению осей координат. Для проведения вращения нужно обязательно указать число факторов, а затем выбрать метод вращения и ответить на вопрос “Нужна ли нормализация Кайзера?”. Нормализация факторных нагрузок Кайзера выполняется для того, чтобы исключить влияние тех переменных, которые имеют по сравнению с другими переменными значительно большие значения нагрузок общих факторов. После расчета факторных нагрузок производится расчет и распечатка коэффициентов общности и специфичности для каждого фактора и, конечно, матрицы факторных нагрузок после вращения. На этом этапе имеется возможность посмотреть факторную диаграмму переменных в осях “фактор 1 — фактор 2”. После просмотра факторных диаграмм можно еще раз вернуться к выполнению процедуры вращения с новым (большим или меньшим) количеством факторов и опять проанализировать факторные диаграммы. После принятия решения о количестве факторов не забудьте сохранить в файле результатов соответствующую матрицу факторных нагрузок — **F2**. При необходимости любую факторную диаграмму можно распечатать на принтере или сохранить рисунок в виде файла.

Работа в системе “SPSS”. После вызова программы из *Windows* так же, как и при работе в “Stadia”, вы попадаете в электронную таблицу (окно редактора данных) и сразу же можете вводить данные в первую переменную (var00001). Если данные уже набраны в виде ASCII-файла, то их можно импортировать в SPSS (меню: **File**, подменю: **Read ASCII Data**). В случае импорта данных следует указать путь к файлу данных и его имя, а также выбрать *тип формата* данных —

Freefield. Далее, нажав на кнопку **Define,** вы переходите *в режим определения переменных*, в котором необходимо каждой переменной (их столько, сколько столбцов в вашем файле данных) присвоить имя — в окошке **Name**, и определить ее тип — **Numeric**. Ввод каждой переменной в общий список анализируемых переменных (Defined Variables) осуществляется нажатием клавиши со стрелкой. После окончания определения *всех* переменных нажмите на клавишу **OK**. SPSS автоматически перейдет в окно редактора данных и осуществит ввод вашего ASCII-файла.

Переход к процедуре факторного анализа осуществляется следующим образом: меню — **Statistics** , подменю — **Data Reduction**, а в нем — **Factor...** После вызова процедуры ФА в правом окне выделите мышкой нужные переменные и перенесите их в окно **Variables**, нажав на кнопку со стрелкой.

Следующий важный этап работы — *выбор параметров* (опций) работы процедуры ФА. Первая группа параметров — расчет необходимых коэффициентов описательной статистики (**Descriptives**). В данном разделе стоит заказать расчет следующих показателей: Univariate descriptives (средние и стандартные отклонения для каждой переменной), Significance level (оценки достоверности получаемых коэффициентов корреляции), а также KMO and Bartlett‘s test of sphericity (соответственно, мера адекватности выборки Кайзера—Мейера—Олкина и коэффициент Бартлета).

Далее выбирают конкретный метод факторизации корреляционной матрицы — **Extraction**. В данном разделе сделайте следующий выбор: 1) в качестве метода укажите — Principal components (метод главных компонент); 2) в подразделе Extract (сколько факторов выделять) можно либо отметить критическую величину собственного значения фактора (Eigenvalues over), например: не меньше 1, либо задать некоторое ожидаемое число факторов (Number of factors); 3) в подразделе Display (какие результаты показывать) выберите пункт Scree plot, чтобы увидеть график изменения собственных значений.

После этого следует выбрать метод вращения осей координат — раздел **Rotation**. Выберите **Varimax,** а также закажите для вывода результатов ФА: Rotated solution (распечатка матрицы факторных нагрузок после вращения) и Loading plots (построение факторных диаграмм).

В разделах Scores и Options все параметры установлены оптимальным образом, поэтому никаких изменений делать не стоит. После установки всех параметров (в каждом разделе не забудьте нажимать кнопку **Continue !**) для начала выполнения процедуры ФА следует нажать кнопку **OK**.

Все текстовые результаты заносятся в окно Output, и их можно просмотреть, используя кнопки скролинга по вертикали (↑ и ↓). Графические результаты ФА находятся в окне **Chart Carusel**, куда можно попасть из головного меню **(Window)** или непосредственно щелкнуть мышью на соответствующей пиктограмме внизу экрана.

Литература

1. *Благуш П.* Факторный анализ с обобщениями. М.: Финансы и статистика, 1989. с. 248.

2. *Иберла К.* Факторный анализ. . М.: Статистика, 1980. 398 с.

3. *Ким Дж.-О., Мьюллер Ч.У.* Факторный анализ: статистические методы и практические вопросы // Факторный, дискриминантный и кластерный анализ. М.: Финансы и статистика, 1989. С. 5 — 77.

4. *Окунь Я.* Факторный анализ. М.: Статистика, 1974. 200 с.

5. *Харман Г.* Современный факторный анализ. . М.: Статистика, 1972. 486 с.

6. SPSS. SPSS Professional Statistics 6.1. Chapter 2. Factor Analysis. Maria J. Norusis / SPSS Inc. 1994. P. 47—82.

Глава 2. МЕТРИЧЕСКОЕ И НЕМЕТРИЧЕСКОЕ МНОГОМЕРНОЕ ШКАЛИРОВАНИЕ

В отличие от всех ранее разработанных методов анализа многомерных наблюдений, таких как факторный анализ, кластер-анализ и т.д., обсуждавшихся подробно в отечественной психофизической литературе, модель многомерного шкалирования (МШ) известна значительно меньше. Это обстоятельство требует детального изложения общих принципов МШ и тех вычислительных алгоритмов, которые будут использованы в учебных заданиях.

§ 1. Основные положения

В основе модели МШ лежит целый ряд предположений о структуре процессов различения объектов-стимулов. Физически каждый объект-стимул характеризуется множеством признаков, например, “объем”, “форма”, “пространственное положение”, “высота”, “длина” и т.п. Сами признаки могут быть простыми и сложными, многомерными. Например, “высота” и “длина” — одномерные признаки, а “форма” и “положение” — многомерные. Отдельный одномерный признак может служить какой-либо одной размерностью более сложного признака. Так, “высота” геометрической фигуры есть одна из размерностей признака “форма”. Каждый стимул характеризуется определенными значениями или степенью выраженности признака.

Точно так же, как стимул характеризуется набором некоторых физических признаков, перцептивный образ стимула (иногда говорят просто — образ) можно характеризовать набором субъективных признаков. Психофизикам известно, что такому, например, физическому признаку стимула, как интенсивность светового излучения, субъективно соответствует яркость; такому, как вес — тяжесть и т.п. Субъективные признаки, также как и физические, могут быть простыми (одномерными) и сложными (многомерными). Однако физическая размерность стимула и субъективная размерность образа в общем случае не совпадают. МШ основывается на положении, что различение стимулов определяется расхождением по ограниченному числу простых субъективных признаков, которые явно или неявно учитываются при суждениях о различии или сходстве стимулов. Исходя из этого положения и ставится главная задача многомерного шкалирования — найти минимальное число субъективных признаков, определяющих различение стимулов человеком, и вычислить значение признаков, которыми характеризуются данные стимулы.

Такая постановка задачи — выявление системы базисных субъективных признаков стимула независимо от их физических коррелятов — позволяет подойти по-новому и к решению основной психофизической задачи — построению функции, связывающей субъективную шкалу стимулов. В отличие от традиционного подхода, когда для заданного физического признака стимулов строится соответствующая субъективная шкала и определяется связывающая их психофизическая функция, МШ дает возможность для заданного субъективного признака стимула определять его физический коррелят, т.е. брать за основу не физическую, а психологическую характеристику стимула. Такой подход к построению психофизической функции может быть полезным для случаев, когда один субъективный признак определяется системой нескольких физических признаков, или когда изменение одного физического признака ведет к изменению сразу нескольких субъективных признаков.

Решение задачи МШ основано на использовании понятия психологического пространства, точки которого представляют исходные стимулы. Аналогично геометрическим представлениям вводится система координат, число которых определяется числом простых субъективных признаков. Это число задает размерность психологического пространства. Оси координат представляют собой шкалы соответствующих субъективных признаков, и положение точек-стимулов в пространстве задано шкальными значениями признаков. Число субъективных шкал и шкальные значения стимулов характеризуют пространственную модель МШ.

Следующее положение, которое также лежит в основании МШ, касается суждений о сходстве или различии между стимулами. Эти суждения предполагаются связанными с положением точек-стимулов в пространстве, поскольку чем более сходны между собой стимулы, тем ближе друг к другу в пространстве представляющие их точки, и наоборот, увеличение воспринимаемого различия между стимулами означает большее пространственное разделение соответствующих точек.

Иначе говоря, предполагается, что расстояние между точками в пространстве есть некоторая функция от субъективного сходства или различия. Метрическая задача МШ заключается в том, чтобы через получаемые суждения о сходстве или различии между стимулами определить расстояния между точками. Решение задачи состоит в построении модели субъективного расстояния в психологическом пространстве.

Формально общая задача МШ выражается следующим образом. По заданной симметричной матрице различий между стимулами

$$D = \left\{ \begin{matrix} D_{11} \ldots D_{1n} \\ \ldots\ldots\ldots\ldots\ldots \\ D_{n1} \ldots D_{nn} \end{matrix} \right\} \qquad (1)$$

нужно построить метрическую и пространственную модели стимулов, т.е. определить размерность пространства и координаты точек-стимулов в этом пространстве

$$X = \left\{ \begin{matrix} X_{11} \ldots X_{1n} \\ \ldots\ldots\ldots\ldots\ldots \\ X_{n1} \ldots X_{nn} \end{matrix} \right\} \qquad (2)$$

таким образом, чтобы матрица расстояний, вычисленных между точками на основании метрической модели расстояния

$$d = \begin{Bmatrix} d_{11} \ldots d_{1n} \\ \ldots\ldots\ldots\ldots \\ d_{n1} \ldots d_{nn} \end{Bmatrix}, \qquad (3)$$

была бы в смысле некоторого критерия возможно более близка к исходной матрице различий D (Терехина, 1986).

В МШ существуют два подхода к решению общей задачи — *метрический и неметрический*. В метрическом МШ на первом этапе строится модель субъективного расстояния. Исходные оценки сходств или различий преобразуются таким образом, чтобы числовые значения удовлетворяли аксиомам геометрического расстояния. На втором этапе по матрице абсолютных расстояний рассчитываются координаты точек и определяется размерность пространства. Для неметрического шкалирования существенными являются не абсолютные числовые значения оценок сходства, а только их порядок. Пространственная модель строится прямо по исходным данным о сходствах или различиях, при этом предполагается, что исходные оценки и межточечные расстояния связаны некоторой неизвестной и монотонной зависимостью, т.е. порядок межточечных расстояний должен соответствовать порядку исходных оценок.

Очевидно, что если исходные данные представлены в виде действительной симметричной матрицы порядка n с элементами не равными нулю, то всегда можно получить конфигурацию точек в пространстве размерности (n-1), удовлетворяющую этому условию. Однако если учитывать главную задачу МШ — определение минимальной размерности пространства, то задача построения пространственной модели сразу становится нетривиальной. Это наглядно иллюстрируется теоремой Гуттмана, которая гласит, что элементы действительной симметричной матрицы порядка n могут быть строго монотонны с расстояниями между n точками в действительном евклидовом пространстве размерностью не более, чем (n-2), только в том случае, если элементы матрицы не равны нулю и не совпадают друг с другом.

Иначе говоря, возможность уменьшения размерности при условии сохранения монотонности связана с дополнительными ограничениями, которым должно удовлетворять искомое решение. Последнее, в свою очередь, означает, что исходные данные должны обладать значительной избыточностью, по сравнению с искомым решением. В каком случае это возможно? Конфигурация точек в пространстве определяется $n \times r$ степенями свободы (где n — число точек-стимулов, r — размерность пространства). Исходная матрица различий имеет c^2 степеней свободы. Следовательно, избыточность исходных данных будет зависеть от того, насколько число стимулов n больше, чем размерность r. Чем больше число стимулов по сравнению с размерностью, тем больше избыточность исходной матрицы и тем более определенной оказывается пространственная и метрическая структура данных, вплоть до нахождения единственного решения, если, конечно, такое решение возможно в принципе. Шепард (1966) показал, что при размерности 2 или 3 для метрического решения практически достаточно 10—15 точек-стимулов.

Таким образом, два неметрических условия, на которые ориентируется решение — монотонности и минимальной размерности — могут дать полную метрическую информацию об исходных данных.

Рассмотрим вкратце принципы достижения монотонности и понижения размерности, которые лежат в основе неметрических алгоритмов.

Достижение монотонности. Условие монотонности означает, что порядок межточечных расстояний d_{ij} должен соответствовать порядку межстимульных различий D_{ij}. Для того, чтобы сделать возможным последовательное сравнение двух порядков, различия и расстояния ранжируются в два отдельных ряда от нуля (минимальная величина) до 1 (максимальная величина). Достижение монотонности есть приведение к нулю всех проранжированных разностей (D_{ij} - d_{ij}), т.е.:

$$\sum (D_{ij} - d_{ij}) = 0 . \qquad (4)$$

Положительное значение (D_{ij} - d_{ij}) означает, что порядок расстояния меньше порядка различия, а отрицательное — что больше. Если данная конфигурация точек (полученная каким-либо произвольным способом) не удовлетворяет условию (4), то конфигурация меняется путем сжатия расстояний с большим рангом и растяжения расстояний с меньшим рангом, чем соответствующий ранг различия. С этой целью для каждой i-й точки по линии, соединяющей ее с j-ой точкой, формируется вектор. Направление вектора определяется знаком разности (D_{ij} - d_{ij}).

Если ранг различия больше ранга расстояния, то вектор направлен от точки i к точке j, а при отрицательной разности вектор направлен обратным образом. Длина вектора зависит от величины различия (D_{ij} - d_{ij}). Для каждой точки i формируется (n-1) подобных векторов. Их общее действие можно представить как действие (n-1)-мерного вектора, приложенного к данной точке i. Перемещение всех точек таким образом приводит к новой конфигурации. Понятно, что новая конфигурация не сразу же после первого шага будет удовлетворять условию монотонности, поскольку каждая точка сдвигается по компромиссному направлению. Процедура достижения монотонности носит итеративный характер и может состоять из значительного числа шагов (Шепард, 1962).

§ 2. Исходные данные. Матрица сходств и различий

Для МШ существенным является определенная организация исходного экспериментального материала в так называемую матрицу сходств. Элементом матрицы (S_{ij}) является некоторая мера сходства между парой стимулов i и j или обратная ей величина D_{ij} — мера различия.

Оценки различий можно получить от испытуемого разными методами. В каждом случае выбор метода шкалирования различий зависит от конкретных экспериментальных условий. Но существует разделение этих методов на два больших класса, которое зависит только от того, какая модель МШ используется для анализа матрицы различий — метрическая или неметрическая.

Условия, налагаемые на элементы матрицы различий в метрическом МШ, строго соответствуют аксиомам расстояния в геометрическом пространстве:

1. Рефлексивность различия:

$$D_{ij} = 0 \qquad (5)$$

подразумевает, что различие между двумя идентичными стимулами (диагональные элементы матрицы различий) должно равняться нулю.

2. Симметричность различий:

$$D_{ij} = D_{ji} \qquad (6)$$

означает, что оценка различия не должна зависеть от временных и пространственных перестановок стимулов относительно друг друга при оценивании (элементы матрицы различий, симметричные относительно главной диагонали).

3. Аксиома треугольника:

$$D_{ij} + D_{jk} \geq D_{ik} \qquad (7)$$

требует, чтобы суммарное различие между любыми двумя парами из трех данных стимулов было не меньше, чем различие между оставшейся парой стимулов.

В терминах теории измерений это означает, что субъективные оценки различий должны представлять собой величины *на шкале отношений*. Только в этом случае их можно рассматривать непосредственно как расстояния между точками в психологическом пространстве или субъективные расстояния.

Методы для шкалирования психологического расстояния между сложными стимулами в большей части прямо аналогичны методам одномерного шкалирования. Большинство методов вполне могут быть расширены до шкалирования многомерных различий. Однако в каждом случае от испытуемого требуется более сложное суждение. Прямое расширение моделей требует некоторой модификации (Торгерсон, 1958). Эти изменения определяются, во-первых, усложнением стимулов, и, во-вторых, сменой содержания оценочных суждений. В одномерном случае оценка представляет величину стимула на шкале, тогда как в МШ оценивается психологическое расстояние между парами стимулов. Если в ситуации одномерного шкалирования шкала отношений или интервалов строилась для самих оценок стимулов, то теперь эти шкалы строятся для межстимульных различий.

Несколько более слабые ограничения налагает на элементы матрицы различий модель неметрического МШ. В общем случае достаточно, чтобы оценки различий удовлетворяли отношениям, установленным для шкалы порядка. Методы порядкового шкалирования основываются на ясных и простых принципах, которые легко реализуются в большинстве экспериментальных ситуаций.

Например, испытуемому могут быть предъявлены все пары стимулов одновременно и он должен упорядочить их по степени сходства. Иногда порядок различия оценивается в баллах.

В некоторых случаях информацию о сходствах можно получить из данных о смешении стимулов. Информацию о смешениях можно получить на основе идентификации испытуемым предъявляемых стимулов. Тогда в клетку ij матрицы заносится число, равное числу случаев, когда испытуемый идентифицировал стимул i как j. Частота случаев идентификации стимула i как j может служить мерой их сходства. Испытуемому можно предложить упорядочить все стимулы в один ряд. Такое упорядочивание производится по отношению к каждому стимулу. Сходство двух стимулов оценивается по частоте их попадания в соседние участки ряда.

Пристального внимания заслуживает вариант, в котором предлагается упростить работу испытуемого, заменив задачу оценивания попарных различий более простой задачей классификации стимулов. Пусть имеется множество многомерных стимулов (цвета, шрифты, вкусовые качества пищевых продуктов, геометрические фигуры и т.п.). Для данного множества стимулов $<N>$ выбирается произвольный набор классов $<\mathrm{k}>$ (категорий, наименований) так, чтобы каждый стимул всегда можно было бы отнести по крайней мере к какому-нибудь одному классу. Набор классов должен исчерпывать классификацию стимулов. Например, для множества вкусовых качеств пищевых продуктов можно предложить набор из четырех основных классов (кислый, сладкий, горький, соленый). Классификация заключается в отнесении каждого данного стимула к одному или нескольким классам. Причем, если стимул относится к одному классу, например, “кислый” для вкуса, то класс заполняется полным весом стимула, или единицей, если же стимул относится сразу к двум классам, например, “кисло-сладкий”, то каждому классу приписывается по половине веса стимула. Если имеет значение место класса в названии, то тому классу, который ставится на первое место, надо приписывать больше веса. Процедура распределения весов стимулов при классификации может быть самой различной, необходимо лишь, чтобы сохранялось порядковое соответствие между распределением весов по классам и предпочтением при классификации стимулов.

В результате классификации стимулов по данному набору классов строится матрица E_{ij}, в которой строка определяется номером стимула S_i - S_n, а столбец указывает класс ($A_i + A_j$). Элементом матрицы E_{ij} является число, показывающее вес стимула S_i по классу A_j, просуммированный по числу предъявлений. Каждая строка матрицы представляет собой вектор $\overline{R}_i$, компонентами которого служат элементы строки $E_{il} \div E_{ik}$. Все строки образуют векторное пространство реакций размерности *k* (по числу классов). В этом пространстве вводится некоторая мера различия между векторами, и тогда попарные различия всех векторов дадут матрицу субъективных различий между стимулами. Полученная таким образом матрица различий вводит данные в систему МШ.

Такая процедура успешно применялась Шепардом и Кэрролом (1966) и Измайловым (1979) к данным называния цветов Бойтона и Гордона (1965) для построения пространства цветоразличения.

§ 3. Построение пространственной модели стимулов

Как уже было сказано, построение психологического пространства предполагает решение двух самостоятельных задач: определения минимального числа осей, необходимых и достаточных для описания структуры межстимульных различий, и вычисления числовых значений, определяющих положение каждого стимула относительно базисных осей координат.

1. Определение базисной размерности.

Определение достаточного числа измерений основано на выборе некоторого критерия, по которому оценивается расхождение между исходной матрицей данных и вычисленными межточечными расстояниями. В идеальном случае это расхождение должно равняться нулю, но в эмпирических данных всегда присутствуют случайные ошибки — шум, величина которого чаще всего неизвестна, поэтому на практике критерий выбирается не нулевой, но достаточно небольшой.

Например, Торгерсон (1958) предлагает следующий метод для определения минимальной размерности. Вычисляется центрированная матрица скалярных произведений между стимулами. Характеристические корни этой матрицы упорядочиваются по величине. Размерность определяется по числу собственных векторов, соответствующих наибольшим характеристическим корням, так, чтобы разброс полученных координат вносил достаточно большой вклад в дисперсию (75—96%). Остальная часть дисперсии рассматривается как следствие случайных ошибок.

Метод определения минимального числа измерений в ходе построения пространственной модели впервые был предложен Шепардом (1962). Он основан на общем принципе понижения размерности, который представляет собой растяжение больших и сжатие маленьких расстояний. Действительно, чтобы поместить, например, треугольник в одномерном пространстве, не нарушая условия монотонности, необходимо сжать его меньшие стороны и растянуть большую. Процедура понижения размерности, так же как и достижение монотонности, основана на формировании множества векторов для каждой точки i, которые должны сжимать маленькие расстояния и растягивать большие. Критерием разделения расстояний на маленькие и большие служит среднее арифметическое расстояний. Поскольку процедура понижения размерности ориентирована на выполнение условия полной монотонности по отношению к различиям, то вместо рангов расстояний, которые на данном шаге итерации не обязательно удовлетворяют условию монотонности, лучше брать ранг самих различий. Тогда вектор, формирующийся для точки i по отношению к точке j, будет определять направление вектора (от точки i к точке j или наоборот), а величина разности (D_{ij} -D) будет определять длину вектора. Сформированные таким образом (n-1) векторы для данной точки i также рассматриваются как действующие аналогично (n-1)-мерному вектору. Как и в ходе достижения монотонности, на каждом шаге итерации меняется положение всех n точек.

При использовании подобных формальных критериев полезно учитывать, что качество аппроксимации исходных данных построенным пространством тем выше, чем больше выбранное число измерений (Спенс, 1972). При увеличении размерности величина ошибки монотонно убывает (рис. 1), поэтому предпочтительнее такое число осей r, при котором эта функция становится достаточно пологой.

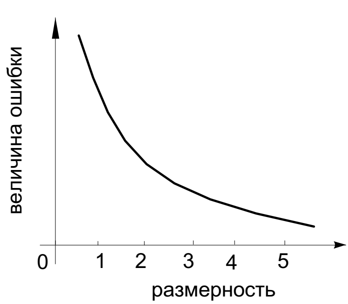


Рис. 1. Примерный вид функции, связывающей величину ошибки с числом измерений пространства.

Формальные критерии определения размерности имеют довольно приблизительную ценность, поскольку в каждом случае выбор критерия оказывается достаточно произвольным. Более важными являются другие критерии, которые основаны на хорошей содержательной интерпретации полученного решения. *Содержательная интерпретация* есть конечный результат производимого анализа, и в любом случае именно она определяет и значимость построенного пространства, и правильность выбора размерности.

Поэтому некоторые авторы (Крускал, 1964) предлагают производить отображения отдельно в одно-, двух-, трех- и т.д. -мерные пространства, строить там оптимальные конфигурации точек и затем выбрать из них такую, которая с точки зрения содержательной интерпретации даст наилучшее решение. Для хорошей интерпретации существенно правильное направление осей координат. В некоторых случаях (Виш и Кэррол, 1974) направление осей координат выбирается в ходе самого алгоритма построения пространственной модели, но в большинстве алгоритмов МШ оси координат имеют произвольное направление, поэтому для облегчения содержательной интерпретации используют *вращение пространства* с тем, чтобы получить оси, связанные с определенными группами стимулов. Аналогичным вспомогательным средством является и метод приведения к главным осям (Терехина, 1974).

Обычно только небольшое число осей получает удовлетворительную интерпретацию, остальные измерения чаще всего являются следствием экспериментального шума. Некоторые измерения могут быть связаны с отдельным подмножеством стимулов, или с данным типом испытуемых, поэтому большой разброс, полученный по данной размерности, еще не означает ее общей важности. Из этого следует очень важный вывод, что окончательное решение не может быть основано на результатах отдельных экспериментов, а необходимо исследование независимых групп данных с привлечением различных методов МШ.

2. Вычисление координат.

К настоящему времени для вычисления координат точек в психологическом пространстве различными авторaми разработано большое количество разнообразных алгоритмов. В данной работе рассматриваются только три из них, которые непосредственно использовались для анализа экспериментальных данных.

Метод ортогональных проекций. Одним из наиболее простых метрических методов МШ является метод ортогональных проекций (Орлочи, 1967; Соколов и др., 1975). Суть его заключается в следующем.

Если есть множество точек, заданных расстояниями между ними, то мы можем максимальное из этих расстояний принять за первую ось (X_1). Точки, заданные этим максимальным расстоянием, обозначим как 1 и 2, затем точку 1 на этой оси примем за начало оси X_1 и спроектируем ортогонально все остальные точки на ось X_1. Тогда точка 1 имеет координату $X_{11} = 0$, а точка 2 — координату $X_{12} = d_{12}$.

Величина проекции для каждой точки i (кроме первых двух точек) вычисляется по известной геометрической формуле:

$$x_{ij} = \frac{d_{1i}^{\ 2} + d_{12}^{\ 2} - d_{2i}^{\ 2}}{2d_{12}}. \qquad (7)$$

Далее легко вычисляются расстояния от каждой точки до оси X_1 по формуле:

$$h_i = (d_{1i}^{\ 2} - x_{1i}^{\ 2})^{1/2}. \qquad (8)$$

Если все h_i=0, то очевидно, что все точки лежат на оси X_1, т.е. пространство данных точек одномерно. Если некоторые h_i>0, то из них выбирается максимальное (h_{max}) и принимается за ось X_2, то есть h_{max}=X_{23}, а точка пересечения h_{max} с осью X_1 есть начало оси X_2 (рис.2).

Затем все остальные точки ортогонально проецируются на ось X_2 и величина проекции вычисляется по формулам:

$$x_{2i} = \frac{x_{23}^{\ 2} + U_i^{\ 2} - d_{3i}^{\ 2}}{2x_{23}}, \qquad (9)$$

где $U_i^{\ 2} = h_i^{\ 2} + (x_{1i} - x_{13})^2.$ (10)

Объединив эти формулы, получим:

$$x_{2i} = \frac{d_{1i}{}^2 + x_{23}{}^2 + (x_{1i} - x_{13})^2 - x_{1i}{}^2 - d_{3i}{}^2}{2x_{23}}. \quad (11)$$

Расстояние от каждой точки до плоскости X_1X_2 определится теперь как

$$q_i = (d_{1i}{}^2 - x_{1i}{}^2 - x_{2i}{}^2)^{1/2}. \quad (12)$$

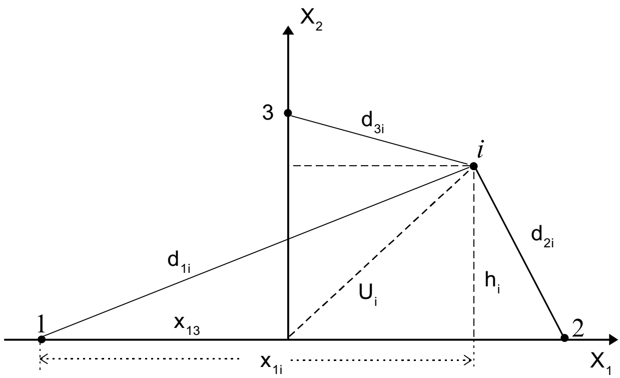


Рис. 2. Определение второй координаты методом ортогональных проекций

И в этом случае, если все $q_i=0$, то, следовательно, все точки лежат на плоскости X_1X_2 и выделение следующей оси пространства не имеет смысла.

Далее, если q_i отличны он нуля, то выбирается точка с максимальным значением — q_{max} и принимается за четвертую точку, и тогда через эту точку 4 будет проходить ось X_3. Далее вычисляются проекции всех остальных точек на эту ось:

$$x_{3i} = \frac{d_{1i}{}^2 + x_{34}{}^2 + (x_{1i} - x_{14})^2 + (x_{2i} - x_{24})^2 - x_{1i}{}^2 - x_{2i}{}^2 - d_{4i}{}^2}{2x_{34}}. \quad (13)$$

Точка, наиболее удаленная от гиперплоскости в пространстве размерности r, ищется из условия:

$$q_i = \max[d_{1i}^{\ 2} - \sum_{k=1}^{r-1} x_{ki}^2] \ . \qquad (14)$$

Процедура продолжается до тех пор, пока сумма всех проекций на k-ю ось не окажется меньше некоторого наперед заданного критерия. Например, эффективность решения можно определять отношением:

$$\frac{\sum_{k} \sum_{i<j} (x_{ki} - x_{kj})^2}{\sum_{i<j} d_{ij}^2} \ . \qquad (15)$$

Обычно ограничиваются таким количеством осей, которое дает разброс, исчерпывающий до 70-90% дисперсии.

Число полученных осей рассматривается как минимальная размерность субъективного пространства, необходимая, чтобы удовлетворялась совместимость всех межточечных расстояний. Простота этого метода делает его удобным для применения к данным, структура которых имеет линейный характер. Однако получающаяся картина существенно зависит от первоначально взятых расстояний, т.е. решение оказывается зависимым от зашумленности исходных данных. Необходимо также отметить, что результирующее пространство определяется всего по нескольким точкам, и поэтому отдельные изолированные точки могут полностью определить решение задачи (Аустин, Орлочи, 1966). Очевидно, что метод, в котором пространство определяется разбросом всех точек, будет иметь более общий характер. Именно такой метод был предложен Торгерсоном (1952, 1958).

Метод Торгерсона. Метод метрического МШ, описанный в работах Торгерсона (1952, 1958), свободен от большинства недостатков метода ортогональных проекций и дает решение, независимое от начального этапа вычислений. Он основан на процедурах аппроксимации исходной матрицы матрицей меньшего ранга (Янг, Хаусхольдер, 1938).

Пусть d_{ij}, d_{ik} и d_{jk} — симметричные расстояния между тремя точками i, j и k. Рассмотрим симметричную матрицу B_i^* с элементами b^*_{ij} и размерностью (n-1)(n-1), где:

$$b^*_{ij} = \frac{d_{ij}^2 + d_{ik}^2 - d_{jk}^2}{2}, j,k \neq 1. \tag{16}$$

Элемент b^*_{ij} представляет собой скалярное произведение векторов от точки i к точкам j и k. Это легко показать с помощью закона косинуса, где для любых трех точек

$$d_{jk} = d_{ij}^2 + d_{ik}^2 - 2d_{ij}d_{jk}\cos\varphi, \text{ откуда:} \tag{17}$$

$$d_{ij}d_{ik}\cos\varphi = \frac{d_{ij}^2 + d_{ik}^2 - d_{jk}^2}{2}. \tag{18}$$

Из уравнений (16) и (18) следует, что $b^*_{ij} = d_{ij}d_{ik}\cos\varphi$, т.е. скалярному произведению векторов из точки i к точкам j и k. Любая из n точек может быть взята как точка i. Таким образом существуют n матриц B*, из которых каждая может быть взята как данная матрица скалярных произведений.

Теоремы Янга и Хаусхольдера показывают, как из матрицы скалярных произведений векторов, начинающихся в точке i, получить информацию о том, возможно ли разместить исходную совокупность точек в вещественном евклидовом пространстве, и если возможно, то какова его минимальная размерность и чему равны координаты точек на этих осях.

Теоремы Янга и Хаусхольдера относятся к любой B_i^* матрице:

1. Если матрица B_i^* положительно полуопределена, расстояния между стимулами могут рассматриваться как расстояния между точками, лежащими в действительном евклидовом пространстве. В терминах характеристических корней или собственных значений матрицы B_i^* это означает, что точки могут рассматриваться лежащими в действительном евклидовом пространстве, если все корни или положительны, или равны 0. Отрицательные характеристические корни предполагают мнимые пространства.

2. Ранг любой положительной полуопределенной матрицы равняется размерности множества точек. Количество положительных значений равняется числу осей, необходимых для описания взаимных межточечных расстояний. Для данного набора стимулов матрица B_i^* будет иметь один и тот же ранг, независимо от того, какой стимул выбран как начало.

3. Любая положительная полуопределенная матрица B_i^* может быть факторизована для получения матрицы *X*, где:

$$B_i^* = X \cdot X' . \qquad (19)$$

Если ранг матрицы B_i^* равен r, где r ≥ (n-1), тогда матрица *X* является прямоугольной матрицей (n-1)×r, элементы которой есть проекция точек на r-ортогональные оси с началом в i-ой точке r-мерного евклидова пространства. Допуская, что для выбора стимулов даны межточечные расстояния (не содержащие случайных ошибок), а матрица B_i^* была построена в заданном начале, различные методики для факторизации матрицы B_i^* дадут различные матрицы *X*, которые, однако, будут связаны ортогональным вращением осей. Матрицы B_i^*, построенные посредством использования различных точек как начала расчета, дадут соответствующие матрицы *X*, которые не отличаются друг от друга с точностью до переноса и вращения осей.

Три теоремы, приведенные выше, определяют решение проблемы пространственной модели стимулов, когда заданы правильные межточечные расстояния. Первая теорема определяет, могут ли стимулы быть представлены точками действительного евклидова пространства. Вторая теорема дает критерии для определения минимальной размерности пространства. Третья теорема дает метод для получения проекций (шкальных оценок) на произвольном наборе осей пространства.

Однако на практике межточечные расстояния всегда даны нам с ошибками. Когда используются ошибочные оценки, то каждая точка будет отчасти ошибочной. Следовательно, в случае допущения, что истинная размерность значительно меньше, чем число стимулов, каждая матрица B_i^* после факторизации даст результаты, которые более или менее различаются из-за ориентации осей и положения начала. При установлении начала в определенной точке возникает неявное предположение, что эта определенная точка является безошибочной. Таким образом мы сталкиваемся с проблемой выбора между n различными факторными матрицами, которые могут быть получены из данных. Одно из решений этой проблемы состоит в том, чтобы поместить начало координат не в какой-либо точке, а в центре тяжести всех точек-стимулов (Торгерсон, 1952, 1958, 1972). Эта процедура дает единственное решение и стремится взаимно компенсировать случайные ошибки для каждой отдельной точки. Опыт показывает, что помещение начала координат в центре тяжести всех точек приводит к меньшим ошибкам, чем помещение его в какую-либо произвольную точку. Рассмотрим процедуру для получения матрицы скалярных произведений векторов с началом в центре тяжести всех точек.

Пусть B_i^* — матрица скалярных произведений размерностью (n-1)(n-1) с центром в точке i. Ее элемент:

$$b^*_{jk} = \sum_e x_{je} x_{ke} \tag{20}$$

Мы будем рассматривать B_i^* как матрицу размерностью (n·n) с i-й строкой и j-ым столбцом, составленными из нулевых элементов. Таким же образом матрицу *X* можно рассматривать как матрицу размерностью (n·r) с i-й строкой, составленной из нулевых элементов.

Наша задача — перенести оси координат из начала в точке i в начало, которое будет центром тяжести для всех точек.

Пусть X^0 — есть искомая матрица проекций точек j на ось e^0 нескольких координатных систем с началом в центре тяжести n точек. Тогда:

$$x_{je^0} = x_{je} - c_{e^0} \quad , \tag{21}$$

где $c_e = \frac{1}{n}\sum_{j}^{n} x_{je} = \frac{1}{n}\sum_{k}^{n} x_{ke}$ (22)

и равно средней проекции точек на ось *e* (т.е. проекции центроида на ось *e*),

$$B^* = X^0 X^{0\prime} \quad (23)$$

и

$$b_{jk}^* = \sum_{e}^{r} x_{je^0} \cdot x_{ke^0} \,. \quad (24)$$

Подставляя уравнение (21) в уравнение (24), получим:

$$b_{jk}^* = \sum_{e}^{r} (x_{je} - c_e)(x_{ke} - c_e)\,. \quad (25)$$

Из определения c_e видно, что

$$b_{jk}^* = \sum_{e}^{r} x_{je} \cdot x_{ke^0} \quad (26)$$

но

$$\sum_{j}^{n} b_{jk}^* = \sum_{e}^{r} x_{je} \sum_{k}^{n} x_{ke} = \sum_{e}^{r} x_{ke} \sum_{n}^{n} x_{je} \quad (27)$$

и

$$\sum_{k}^{n}\sum_{j}^{n} b_{jk} = \sum_{e}^{r} (\sum_{j}^{n} x_{je})^2 \,. \quad (28)$$

После подстановок получим:

$$b_{jk}^* = b_{jk} - \frac{1}{n}\sum_{j}^{n} b_{jk} - \frac{1}{n}\sum_{k}^{n} b_{jk} + \frac{1}{n^2}\sum_{j}^{n}\ \sum_{k}^{n} b_{jk}\,. \quad (29)$$

Подстановка b_{jk} из уравнения (16) в уравнение (29) дает:

$$b^*_{jk} = b_{jk} - \frac{1}{n}\sum_{j}^{n} b_{jk} - \frac{1}{n}\sum_{k}^{n} b_{jk} + \frac{1}{n^2}\sum_{j}^{n}\ \sum_{k}^{n} b_{jk}\ . \qquad (30)$$

Суммирование каждого выражения в отдельности и упрощение приводит к:

$$b^*_{jk} = \frac{1}{2}(d_{ij}^{\ 2} + d^2{}_{ik} - d^2{}_{jk}) - \frac{1}{2n}\sum_{j}^{n}(d^2{}_{ij} + d^2{}_{ik} - d^2{}_{jk}) -$$
$$- \frac{1}{2n}\sum_{k}^{n}(d_{ij}^2 + d_{ik}^2 - d_{jk}^2) + \frac{1}{2n^2}\sum_{j}^{n}\sum_{k}^{n}(d_{ij}^2 + d_{ik}^2 - d_{jk}^2)\ . \qquad (31)$$

Уравнение (31) дает стандартный прямой метод, чтобы из межточечных расстояний вычислить матрицу B^* скалярных произведений с началом координат в центре тяжести всех точек. Получив оценки субъективных расстояний между всеми парами стимулов, уравнением (31) пользуются для вычисления матрицы B^*. Затем эта матрица факторизуется посредством любого из обычных методов факторизации для получения проекций стимулов на r ортогональных осей пространства, проходящих через центр тяжести всех объектов. Ориентация осей зависит как от конфигурации, так и от выбранного способа факторизации (Торгерсон, 1958). Матрица B^* факторизуется в общем аналогично корреляционной или ковариационной матрице в методе главных компонент (Айвазян и др., 1974). По сравнению с другими ортогональными преобразованиями, преобразование к главным компонентам искажает структуру исходных данных наименьшим образом, поскольку всякий набор из данного числа главных компонент характеризует максимальный разброс точек, спроектированных в пространство этих главных компонент. Выбор числа главных компонент определяется величиной суммарной дисперсии, которую необходимо исчерпать в том или ином решении. На практике ограничиваются теми главными компонентами, которым соответствуют наибольшие характеристические корни, а все остальные компоненты отвергаются как незначительные.

Таким образом, метод Торгерсона дает возможность построить оптимальную пространственную модель стимулов в том смысле, что полученное решение не зависит от случайных экспериментальных ошибок, поскольку оно определяется структурой сразу всех стимулов. Однако необходимо учитывать, что пространственное представление стимулов, определенное по нескольким максимальным характеристическим корням матрицы скалярных произведений, может оказаться непригодным, если, например, истинная структура стимулов имеет локальные нелинейные цикличности (Терехина, 1977).

Алгоритм Янга-Торгерсона. Построение пространственной модели производится в два последовательных этапа. На первом этапе исходная матрица различий анализируется метрическим методом Торгерсона. По числу наибольших характеристических корней определяется размерность пространства, и таким образом формируется исходная конфигурация для n точек, между которыми вычисляются n(n-1)/2 расстояний.

На втором этапе данная конфигурация проверяется на выполнение условия монотонности. Для этого строится диаграмма монотонности. Она представляет собой график, осью абсцисс которого служат межточечные расстояния, а осью ординат — исходные различия. Каждой паре точек-стимулов (i,j) на этой диаграмме будет соответствовать точка с абсциссой d_{ij} и ординатой D_{ij}. Условие монотонности означает, что от начальной точки графика каждая последующая точка должна располагаться только правее или выше предыдущей, и никогда не может быть ниже или левее. Если, следуя этому правилу, соединить последовательно все точки отрезками, то получится график, характеризующий монотонность связи между межточечными расстояниями и исходными различиями. Очевидно, что если для каких-либо пар точек-стимулов (i,j) монотонность не выполняется, то точки, представляющие их на диаграмме монотонности, не попадут на построенный график, а будут левее или ниже его. Для каждой выпавшей из графика точки можно вычислить ее отклонение от графика по оси абсцисс расстояний (по оси ординат это отклонение измерять не нужно, поскольку порядок различий задан как исходный) и сумма этих отклонений

$$\sum(d_{ij}-\hat{d}_{ij})^2\ , \qquad (32)$$

$\hat{d}_{ij}$ есть значение “правильной” абсциссы, для точки, выпавшей из графика монотонности) покажет степень несоответствия данной диаграммы условию полной монотонности.

Данный метод определения количественной меры достижения монотонности был предложен Крускалом (1964) и предложенная им мера была названа *стрессом*. Для достижения полной монотонности в алгоритме Янга—Торгерсона определяется матрица минимизирующая выражение (33). Количественной мерой достижения монотонности служит мера, названная авторами индексом адекватности:

$$I=\frac{1}{2}+\frac{1}{2}\frac{\sum d_{ij}\hat{d}_{ij}}{(\sum d_{ij}^2\cdot\sum\hat{d}_{ij})^{1/2}}\ . \qquad (33)$$

В случае неудовлетворительного значения *I* найденные $\hat{d}_{ij}$ вводятся в исходную матрицу различий, и первый этап повторяется сначала. В обратном случае переходят к следующему этапу — улучшению исходной конфигурации. Для этого каждая координата *X* точки i преобразуется с помощью выражения:

$$\overline{x}=x_i+\frac{p}{n-1}\sum\frac{(d_{ij}-\hat{d}_{ij})}{d_{ij}}(x_k-x_i)\ , \qquad (34)$$

где *p* — коэффициент, определяющий сходимость итеративного процесса.

Новая конфигурация также проверяется на условие достижения монотонности по индексу адекватности *I*. Если

I не получил удовлетворительного значения, то последний этап повторяется. В обратном случае решением являются координаты точек, полученные в последнем шаге итерации.

§4. Построение метрической модели

В ходе построения пространственной модели данных необходимо измерять расстояния между точками-стимулами, чтобы соотносить их с исходными оценками различий. Для измерения расстояний в пространстве вводится метрика. Выбор метрики для психологического пространства также основывается скорее на содержательных аспектах данных, чем на формальных.

Так, Шепард (1964) предлагает условное деление стимулов на два класса в зависимости от их перцептивной целостности. Имеется в виду, что одни стимулы воспринимаются как целостные образования, и обычно сознательно не анализируются, например, цвета, запахи, фонемы и т.п., тогда как другие стимулы явно различаются по несвязанным между собой признакам, как, например, в работе Эттнива (1950), где стимулы — плоские геометрические фигуры — различались по величине, яркости и ориентации. В пространственной модели “не анализируемых” стимулов удобнее использовать евклидову метрику. Инвариантность евклидова расстояния относительно вращения систем координат (изотропность евклидова пространства) соответствует в данном случае такому типу поведения испытуемого, как если бы он оценивал различия между простыми, одномерными объектами. В случае явно “анализируемых” стимулов, когда итоговая оценка составляется как бы из последовательного добавления по очередному признаку, более подходит “sity-block”-метрика.

И метрика “sity-block”, и евклидова метрика являются частными случаями одной общей функции:

$$d_{ij} = [\sum_{k=1}^{r} (x_{ik} - x_{jk})^p]^{1/p} \ , \qquad (35)$$

известной как метрика Минковского. Для случая “sity-block” p=1, а для евклидовой метрики — p=2.

Конечно, выбор метрики определяется не только тем, “анализируемые” стимулы или “не анализируемые”, и не ограничивается двумя приведенными видами метрик. В некоторых работах предлагается решать задачу для нескольких значений p, и затем экспериментатор выбирает наиболее “интерпретируемую” модель расстояния. Например, для пространственной модели цветоразличения Крускел (1964), варьируя в выражении (35) показатель p от 1 до 5, получил, что в данном случае наиболее подходит метрика с p=2.5. В другой работе (Шепард, 1962) было показано, что в пространстве цветоразличения можно принять евклидову метрику, если использовать нелинейную форму соотношения между исходными мерами сходства и межточечными расстояниями.

Такие выводы все-таки носят частный характер, они связаны с конкретными экспериментальными ситуациями, тогда как для более общих выводов, как и в случае построения пространственной модели, необходимы более широкие исследования.

Существенное влияние на вид метрики может оказать инструкция, направляющая внимание испытуемого. Например, в одном из опытов Шепард (1962) предъявлял испытуемому стимулы, представляющие собой окружность с одной радиальной линией. Стимулы различались между собой величиной окружности и углом наклона радиальной линии. Исследования показали, что результаты оценок зависят от того, на что больше обращает внимание испытуемый — на величину окружности или на наклон радиуса. Причем одни испытуемые оценивают стимулы только по одному какому-то признаку и стараются последовательно придерживаться выбранной стратегии, другие стараются учитывать оба признака, а третьи могут переключать внимание с одного признака на другой. Поэтому при проведении исследований следует искать решение отдельно для каждой фиксированной инструкции. При построении общей модели полезно сопоставлять данные, полученные для различных установок внимания.

§5. О развитии моделей многомерного шкалирования

Выше были описаны так называемые *классические* модели метрического и неметрического МШ (Торгерсон, 1952; Шепард, 1962; Крускал, 1964). Их характерная особенность заключается в том, что анализируется лишь одна матрица различий. В тех же случаях, когда исследователь имеет несколько таких матриц, то он вынужден либо анализировать из по отдельности, либо усреднять все данные, сводя их в одну матрицу.

Следующим серьезным вкладом в разработку новой идеологии в МШ (после работ Шепарда и Крускала) была разработка *Мак Ги* (1968) так называемого *реплицирующего МШ* (replicated MDS), распространившего МШ на одновременный анализ более чем одной матрицы сходств. Характерной особенностью этого подхода является то, что она применяет одну и ту же модель евклидовой метрики к нескольким матрицам различий *одновременно*. Основное допущение данного подхода заключается в том, что всем отдельным матрицам данных соответствует одна и та же пространственная конфигурация стимулов. Из этого следует, что, с точностью до случайной ошибки все матрицы одинаковы, и, таким образом, повторяют одна другую. Используя процедуру реплицирующего МШ (например, в системе SPSS), исследователь получает возможность одновременно анализировать несколько отдельных матриц и строить единое субъективное пространство по данным нескольких испытуемых. Хорошим примером использования данного подхода в МШ может служить работа Якобовича (1974), где предпринято исследование развития речи у детей. В его эксперименте детей 5, 7, 9 лет и взрослых (по 15 человек в группе) просили оценить различие между 15 парами частей человеческого тела. Данные по каждой группе в отдельности (15 повторяющихся матриц) обрабатывались реплицирующим МШ.

Другим серьезным продвижением в МШ (после разработки неметрического МШ) по праву считают работы Кэррола и Чанга (1970). Поскольку первоначальная модель Торгерсона не допускала каких-либо индивидуальных различий в процессе оценивания испытуемыми сходства стимулов, а индивидуальные различия представляют для психологов особый интерес, этими авторами была разработана новая модель — *индивидуальное шкалирование* или *взвешенная модель* МШ. В этой обобщенной модели, основанной также на евклидовой метрике, предполагается, что между несколькими исходными матрицами могут быть нелинейные и немонотонные различия. Таким образом, предполагается существование индивидуальных различий гипотетических субъективных пространств испытуемых. Название "взвешенная" модель МШ получила в силу предположения о том, что координаты стимулов для каждого отдельного испытуемого связаны с координатами групповой матрицы некоторыми весовыми коэффициентами. Эти весовые коэффициенты и являются оценками индивидуальных различий.

Модели индивидуальных различий в МШ достаточно широко применяются в психологии, главным образом для изучения индивидуальной специфики оценок сложных стимулов различными людьми. В ряде современных статистических систем представлены "хорошие" реализации взвешенной модели МШ, например, так называемая процедура INDSCAL в системе SPSS.

Литература

Основная

1. *Дейвисон М.* Многомерное шкалирование. М.: Финансы и статистика, 1988.

2. *Парамей Г.В.* Применение многомерного шкалирования в психологических исследованиях // Вестн. Моск. Ун-та. Сер. 14, Психология. 1983. № 2. С. 57—70.

3. *Терехина А.Ю.* Анализ данных методами многомерного шкалирования. М.: Наука, 1986.

Дополнительная

1. *Измайлов Ч.А.* Сферическая модель цветоразличения. М.: Изд-во Моск. ун-та, 1980.

2. *Торгерсон У.С.* Многомерное шкалирование. Теория и метод // Статистическое измерение качественных характеристик. М.: Статистика, 1972.

3. *Шепард Р.* Многомерное шкалирование и неметрические представления. // Нормативные и дескриптивные модели принятия решений. М.: Наука. 1981.

4. Multidimentional Scaling. SPSS Professional Statistics 6.1. Marija J. Norusis / SPSS Inc. 1994. P. 155-222.

Методические указания по выполнению учебного задания по теме: “Многомерное шкалирование“

Задание 1. Построение субъективного пространства восприятия букв русского алфавита

Цель задания. *практическое освоение метода МШ и овладение навыками применения его в психологических исследованиях для измерения сложных субъективных переменных и построения геометрической модели различения сложных стимулов.*

Методика

Аппаратура. Задание выполняется на IBM-совместимом персональном компьютере. Для предъявления сигнала “Внимание” используются головные телефоны, соединенные со звуковым синтезатором персонального компьютера. Для выполнения учебного задания используется универсальная компьютерная программа *parcom.exe*.

Стимуляция. В центре экрана монитора предъявляются пары букв русского алфавита[1]. Длительность предъявления — 1 с.

Процедура опыта. При отработке задания каждый студент выступает сначала в роли испытуемого, а затем обрабатывает собственные данные.

Проба начинается со звукового сигнала “Внимание” (тон частотой 500 Гц и длительностью 200 мс). После

[1] Набор стимулов может варьироваться экспериментатором, для чего нужно внести изменения в файл *stim.dat.*

предъявления стимулов в течение четырехсекундного интервала испытуемый должен оценить по 10-ти балльной шкале степень *различия* букв в паре: 0 — минимальное различие, 9 — максимальное. Оценки фиксируются на цифровой клавиатуре компьютера и подтверждаются нажатием клавиши “Enter”. Всего в опыте используется 12 букв, что составляет n(n-1)/2=66 пар. Каждая пара предъявляется по 4 раза. Предъявления производятся в квази-случайном порядке.

Опыт состоит из тренировочной (10 проб) и основной серий (264 пробы); в середине основной серии — двух-трех минутный перерыв.

Обработка результатов

После окончания опыта каждый студент получает распечатку матрицы субъективных различий 12 × 12, каждый элемент этой матрицы является результатом усреднения 4-х оценок. Эти данные также записываются в обычный ASCII-файл, имя которого соответствует фамилии испытуемого, а расширение — *mds* (например: *ivanov.mds*).

Для дальнейшей обработки результатов следует воспользоваться одной из современных статистических программ, в состав которой входят методы МШ. Ниже будет показано, как обрабатывать результаты в статистических системах *“Stadia”* и *SPSS* .

<u>Работа в системе “Stadia”</u>. После запуска программы (stadia.exe) пользователь попадает в редактор данных и перед ним открываются 2 возможности: либо вручную ввести свои данные в электронную таблицу, либо прочитать файл с матрицей данных с жесткого диска или дискеты (F3). Напомним, что данные должны быть в виде квадратной матрицы субъективных различий; матрица может быть либо полной, когда обе ее половины симметричны относительно отрицательной диагонали, либо данные заносятся лишь в ее нижнюю левую половину, а верхняя может быть заполнена любыми числами, например, нулями. Далее в меню статистических процедур (F9) нужно выбрать пункт “Многомерные методы”, а в этом пункте — подпункт “Шкалирование”.

После окончания этапа факторизации матрицы различий на экране появляются результаты первых расчетов: собственные значения координатных осей и процент объясняемой дисперсии. Смысл этих показателей подробно объяснен в главе 1 настоящего раздела, посвященной факторному анализу. Сразу же после появления на экране первых результатов необходимо сохранить их в отдельном файле для последующего анализа, для чего следует нажать на клавишу F2, выбрать опцию записи данных в файл и задать имя файла данных. В дальнейшем после появления на экране монитора новых результатов и нажатия на F2 они будут записываться в тот же самый файл.

Далее нужно выбрать один из вариантов метода — метрический или неметрический. В данном учебном задании предусматривается знакомство с обоими вариантами. Метрическое МШ представлено алгоритмом Торгерсона, в неметрическое включены 4 варианта, отличающиеся способом расчета "стресса".

В метрическом алгоритме Торгерсона предусмотрен расчет координат стимулов-объектов, исходя из *евклидовой метрики.* В распечатке результатов представлены: 1) координаты стимулов в n-мерном пространстве; 2) четыре различных показателя "стресса", характеризующие "хорошесть" полученного решения с точки зрения соответствия исходной матрицы различий матрице вычисленных расстояний между стимулами.

При обработке матрицы субъективных расстояний по одному из неметрических алгоритмов, следует сначала выбрать одну из метрик (для корректности сравнения с результатами метрического анализа следует также выбрать евклидову метрику), а затем указать ожидаемое количество осей субъективного пространства. При выборе количества осей следует руководствоваться следующими соображениями: 1) начинать лучше с явно избыточного количества, например, с четырех или даже с пяти, а затем, последовательно уменьшая размерность пространства, внимательно анализировать изменение показателей "стресса" и пытаться дать выделенным осям содержательную интерпретацию[1]; 2) в любом случае не следует ограничиваться каким-то одним вариантом (скажем, двумерным, и только потому, что такую модель легче изобразить геометрически), для последующего анализа результатов стоит попробовать несколько решений и получить соответствующие распечатки.

В распечатке результатов представлены те же показатели, что и в метрическом методе — координаты стимулов в пространстве указанной размерности и показатели "стресса".

Работа в системе SPSS. Для тех студентов, кто знает английский язык и ориентируется в операционной среде Windows, можно также рекомендовать статистическую систему *SPSS*, которая нашла достаточно широкое распространение среди наших зарубежных коллег. Эта система не так проста и доступна как "Stadia", но зато предоставляет ряд дополнительных возможностей, которые могут быть весьма полезны психологу в серьезной исследовательской работе.

После вызова программы из Windows так же, как и при работе в "Stadia", вы попадаете в электронную таблицу редактора данных и сразу же можете вводить данные в первую переменную (var00001). Если данные уже набраны в виде ASCII-файла, то их можно импортировать в SPSS (меню: File, подменю: Read ASCII Data). В случае импорта данных следует указать путь к файлу данных и его имя, а также выбрать подходящий тип формата данных — Freefield. Далее, нажав на кнопку Define, вы переходите в режим определения переменных, где необходимо каждой переменной (их столько, сколько столбцов в вашем файле данных) присвоить имя — в окошке Name, и определить ее тип — Numeric. Ввод каждой переменной в общий список анализируемых переменных (Defined Variables) осуществляется нажатием клавишу Add. В случае какой-либо ошибки при вводе имени переменной следует воспользоваться

---

[1] Стратегия исследователя при выборе оптимального количества осей многомерного пространства подробно описана в предыдущей главе.

клавишей Remove, и осуществить ввод заново. После окончания определения всех переменных нажмите на клавишу OK. SPSS автоматически перейдет в окно редактора данных и осуществит ввод вашего ASCII-файла.

Переход к процедуре МШ осуществляется следующим образом: меню — Statistics, подменю — Scale, а в нем — Multidimentional Scaling... После вызова процедуры МШ в правом окне выделите мышкой все переменные и перенесите их в окно Variables, нажав на кнопку со стрелкой.

Следующий важный этап работы — выбор основных параметров процедуры МШ. В первой группе параметров (Distances) указывают, что собой представляют исходные данные. В нашем случае — это квадратно-симметричная матрица различий (square symmetric), поэтому данный параметр, установленный в программе по умолчанию (т.е. исходно), изменять не нужно. Если же в опыте получена ассимметричная матрица, и у исследователя есть основание думать о невыполнении аксиомы симметричности, то следует указать другую форму матрицы данных (Shape) — квадратно-асимметричную (square asymmetric). Кроме того, в данной опции имеется возможность работать не только с матрицей различий, но и получать оценки субъективных различий не прямо (в результате их прямого сравнения), а косвенно — расчетным путем, например, из оценок испытуемыми предъявляемых стимулов по отдельности (Create distances from data). Для выхода из данного пункта и для подтверждения всех сделанных установок нужно нажать на кнопку Continue (продолжить).

Следующий важный шаг — выбор параметров модели МШ (Model...). Главное в данном пункте — это задать уровень проведенных измерений (Level of Mesurement), т.е. указать, с какими данными вы имеете дело — метрическими (Interval, Ratio) или неметрическими (Ordinal). Как было сказано выше, данным, измереннным по шкале порядка (Ordinal), адекватна неметрическая модель МШ, а более “сильным” данным, измеренным по шкалам интервалов (Interval) или отношений (Ratio), соответствует метрическая модель. Кроме того здесь же указывается минимальное и максимальное количество ожидаемых осей многомерного пространства (Dimensions); очевидно, что в процессе поиска наилучшего решения эти два параметра могут меняться. В двух оставшихся пунктах (Scaling model и Conditionality) следует оставить без изменения те параметры, которые заданы по умолчанию, т.е. Euclidean distance и Matrix, соответственно; эти две опции имеют отношение к более сложным моделям МШ, которые в рамках данного учебного пособия не рассматриваются. После установки всех параметров необходимо также нажать на кнопку Continue (продолжить).

В последней группе параметров (Options) заказывают дополнительные возможности вывода результатов МШ и ряд числовых критериев, определяющих нюансы работы расчетного алгоритма МШ. В первой группе (Display) стоит заказать выдачу графика геометрической модели субъективного пространства (Group plots), а также возможность записи в файл результатов информации обо всех выбранных выше параметрах (Model and options summary). Во второй группе опций (Criteria) не стоит делать никаких изменений — пусть все параметры останутся заданными по умолчанию. Заметим лишь, что эти параметры определяют количество итераций вычислительного алгоритма МШ и, в принципе, могут повлиять на точность и оптимальность вычислений координат стимулов. После нажатия на кнопку Continue (продолжить) вы возвращаетесь в основное меню и можете запустить выполнение процедуры ALSCAL МШ, нажав на кнопку OK.

В распечатке результатов (окно Output1) приводится следующая информация (по порядку):

1. Сводка всех выбранных параметров модели и опций.

2. Изменение величины “стресса” (по Янгу) на каждом шаге итерации. Отметим, что в соответствии с заданным в опциях критерием вычисления прекращаются в том случае, когда: а) величина стресса становится меньше заданной величины (Minimum S-stress = 0,00500); б) изменение величины стресса на следующем шаге итерации очень незначительно (Convergence Criterion = 0,00100); в) до-

стигается максимальное число итераций (Maximum Iterations = 30).

3. Величины “стресса” (по Крускалу) и квадрата коэффициента корреляции (RSQ) исходной матрицы субъективных различий и матрицы межстимульных расстояний, вычисленных моделью. Это основные показатели для оценки “хорошести” соответствия расчетов использованной модели МШ исходным данным.

4. Координаты стимулов в n-мерном пространстве. Напомним, что минимальная и максимальная величины n задавались в подменю Dimentions при выборе параметров модели.

5. В самом конце распечатки указывается, какие построены графики.

Нажав на пиктограмму Chart Carusel внизу экрана или выбрав в основном меню SPSS пункт Windows (а в нем подпункт Chart Carusel), имеется возможность проанализировать построенные графики. Основной рисунок — это график расположения стимулов в 2- или 3-мерном эвклидовом пространстве (Derived Stimulus Configuration) или геометрическая модель субъективного пространства, построенная по рассчитанным выше координатам стимулов (см. распечатку результатов, пункт 4). Кроме того полезно также посмотреть и на график соответствия исходных различий (Disparities) рассчитанным моделью расстояниям (Distances). Эта так называемая диаграмма рассеяния (Scatter plot of linear fit) является наглядным графическим представлением качества этого соответствия, и, по сути дела, соответствует указанному выше коэффициенту корреляции (RSQ). Фактически, чем лучше точки ложатся на прямую линию, тем лучше модель воспроизводит исходные данные.

Анализ и обсуждение результатов

В ходе анализа полученных результатов (начнем с метрического варианта решения) нужно рассмотреть проблему *определения минимальной размерности* полученного пространства. Для ее решения необходимо сопоставить величины собственных значений осей и процент объясняемой дисперсии для пространств различной размерности. Значительное понижение величины собственного значения при переходе от n-мерного решения к (n+1)-мерному свидетельствует о достаточности n измерений и избыточности (n+1)-й координаты. Близкие значения для обоих решений скорее свидетельствуют в пользу необходимости учитывать и (n+1)-ю координату.

"Важность" каждой следующей оси анализируемого пространства оценивают также и по величине ее вклада в общий процент объясняемой дисперсии. Строго говоря, какой-либо определенный рецепт о критической величине суммарного процента объясняемой дисперсии дать достаточно трудно. Здесь лучше полагаться на здравый смысл и опыт. Тем не менее, будет уместно сделать два замечания: во-первых, явно не стоит добиваться (и поэтому ожидать) очень высокого суммарного процента, поскольку в измерениях всегда присутствует известная доля экспериментального "шума"; во-вторых, опыт многих психофизиков показывает, что в пилотажных работах *70—75%* объясняемой дисперсии считается очень хорошим показателем, и даже в хорошо спланированных исследованиях он не часто превышает *90%*.

Вывод о минимальной размерности должен быть согласован с содержательной интерпретацией осей координат, которая составляет вторую важную проблему анализа результатов. На этой стадии анализа результатов необходимо построить 2- и 3-мерные графические модели субъективного пространства и очень внимательно проанализировать взаимное расположение точек-стимулов. При невозможности построить 3-мерные диаграммы вручную, можно ограничиться анализом 2-мерных графиков (оси: 1-2, 1-3, 2-3, 1-4 и т.д.) либо воспользоваться любой компьютерной программой, обеспечивающей 3-мерную графику. Это несложно сделать в программе "Stadia", для чего введите матрицу координат стимулов в редактор данных и воспользуйтесь функцией "Рисунок" (F6) из основного меню, затем выберите в меню "Тип графика данных" пункт 9 — "3-мерный данных" и укажите номера изображаемых переменных.

По графикам необходимо определить *графические характеристики* букв, связанные с каждой осью полученного пространства, и дать им по возможности простую и однозначную интерпретацию.

После определения минимальной размерности субъективного пространства следует обратиться к результатам неметрического шкалирования и обработать матрицу субъективных различий одним из предложенных методов, выбрав евклидову метрику и задав определенное выше число осей. После построения графиков результатов неметрического решения их нужно сравнить с аналогичными графиками результатов метрического решения и проанализировать сходство и различие.