



УДК 159.9.07

## **О возможности построения шкалы интервалов в психологических измерениях**

А. И. Худяков

*Российский государственный педагогический университет  
им. А. И. Герцена, г. Санкт-Петербург  
E-mail: [haipsy@yandex.ru](mailto:haipsy@yandex.ru)*

**Аннотация.** Основная проблема психологических измерений заключается в разрыве между психологическим анализом процесса измерения и соответствующими математическими моделями. Математическая модель любого процесса будет действенной, если имеется априорное представление о сути процесса. Когда мы, например, используем вероятность, тем более распределенную по нормальному закону, мы не решаем проблему, а просто обходим ее острые углы. По сути, в этом случае мы обращаемся с объектом исследования как с черным ящиком. По результатам прямых методов психологического измерения стимулов, не имеющих физической метрики, можно построить метрические шкалы – интервалов и отношений. Структура этих шкал снимает вопрос о конкретной единице измерения, так как соответствующие допустимые преобразования позволяют работать с условными единицами.

**Ключевые слова:** психологическое измерение, психофизическое шкалирование, шкалы интервалов, эмпирические процедуры, вероятностные модели, метод парных сравнений, непосредственная количественная оценка, непосредственная количественная оценка субъективных интервалов.

В психологии возможность построения шкал интервалов до сих пор вызывает споры. Одни категорически утверждают, что это невозможно в принципе, другие не менее категорично утверждают, что это возможно, но требуют более корректного обоснования выбора психолого-математической модели.

К огромному сожалению, до сих пор в отечественных публикациях встречается очень невнимательное отношение к проблеме измеримости и возможностей применения метрических шкал в психологическом исследовании. Это хорошо просматривается на материале педагогических конструкций, утверждающих равенство расстояний в ранговых шкалах, например, при оценке школьных знаний по пятибалльной шкале. Хотя очевидно, что эта шкала имеет только четыре градации и расстояния между рангами очень далеки от равных. Это легко показать, используя закон сравнительных суждений Л. Л. Терстона и метод парных сравнений. Однако такие изыски недоступны чаще всего авторам многих диссертаций и публикаций, и они вслед за первоначально неправильным положением держатся очень уж упрощенного понимания процесса психологического измерения.

Как и в случае других измерений, результатом психологического измерения является случайная величина. Вид распределения этой случайной величины показывает степень устойчивости результатов измерения. Если распределение равномерно по всему интервалу возможных значений оценок или величина эксцесса очень велика, то можно сделать вывод, что нам не удалось обнаружить предполагаемый фактор, или предложенный критерий оценки неадекватно понят испытуемыми, или выборка испытуемых неоднородна. В последнем случае есть надежда, разбив исходную выборку испытуемых каким-либо способом на однородные подвыборки, получить набор распределений с выраженным максимумом. Наличие максимума распределения показывает, что мы что-то с какой-то точностью (мера вариабельности) измерили – выявили какой-то латентный фактор, но насколько этот фактор соответствует искомому (проблема валидности) по виду распределения, разумеется, решить невозможно.

При отсутствии соответствующих «опорных» физических шкал вид распределения оценок – единственный показатель устойчивости результатов измерения, а проблема интерпретации латентного фактора аналогична проблеме интерпретации в факторном анализе.

Если разные эмпирические процедуры, в основе выполнения которых лежат разные психические системы, дают результаты, «попадающие» в окрестности «истинного» латентного фактора, то можно надеяться на успех проведения измерения, но какое «попадание» ближе к цели, решить практически невозможно [3].

Следовательно, даже меньший разброс результатов, полученных с помощью какой-либо процедуры (большая точность), не дает права утверждать, что эти результаты более валидны и что эта эмпирическая процедура предпочтительна.

Если нет оснований для выбора одного из нескольких конкурирующих методов, то целесообразно применять их все и считать решением то, что есть общего в результатах [4]. Как правило, к результатам, полученным различными эмпирическими процедурами, применяют различные математические модели, что затрудняет сравнение полученных шкал, практически не дает возможности для взаимной проверки данных, полученных разными способами.

Вероятностные модели построения психологических шкал имеют ряд неоспоримых преимуществ. Главное преимущество в том, что введение вероятностной меры позволяет корректно совместить психологические и математические теоретические модели и, как следствие, использовать не только вычислительную, но и прогностическую мощь математики. Но при этом возникает опасность отрыва теоретических, абстрактных построений от реальности, что проявляется в расхождении теоретических прогнозов и данных эксперимента. Существенный недостаток вероятностного подхода обусловлен свойствами самой вероятностной меры. Вероятностная мера вводится на множестве событий и множестве исходов. Оценка конкретного вида этой меры (функция распределения) производится по виду распределения частот в гистограмме. Вид распределения частот оценивается по выборке. При этом отдельная реализация не имеет смысла сама по себе. Имеет смысл только некая совокупность реализаций, попадающих в некий интервал. Так же как в теории вероятности не имеет смысла утверждение «попадание в точку числовой оси с такой-то вероятно-

стью», а имеет смысл утверждение «попадание с такой-то вероятностью в данный интервал». Поэтому действия с отдельными реализациями не корректны в парадигме вероятностной модели, следовательно, невозможен индивидуальный подход. Имеют смысл только данные группы испытуемых, допускающие вероятностные представления, а не данные отдельного индивида.

Кроме того, оценка значений вероятностей по полученным в эксперименте частотам приводит к тому, что при переходе к табличным  $z$ -оценкам проблема погрешности измерения остается в стороне. Причем небольшие выборки, используемые в психологических экспериментах (объемом в 20–30 чел.), приводят к достаточно грубым оценкам вероятностей.

Рассмотрим альтернативный подход, лишенный этих недостатков. Исходным элементом в предлагаемой модели является набор оценок, данных отдельным испытуемым стимульному набору.

В методе парных сравнений оценкой стимула будем считать количество «побед», одержанных этим стимулом в актах сравнения с другими стимулами. В процессе измерения испытуемый распределяет данное количество очков (возможных побед) среди стимулов. Всего возможно  $n(n-1)/2$  актов сравнения, где  $n$  – количество стимулов, следовательно, это количество очков распределяется между стимулами. Максимальное число побед, которое может получить один стимул, равно  $n-1$ , минимальное, соответственно, 0. Чем более в стимуле выражен критерий оценки, тем чаще этот стимул выигрывает в сравнении с другими, тем больше получит очков. В результате применения метода парных сравнений к выборке испытуемых мы получаем трехмерную матрицу  $A_{ijk}$ , где  $i, j = [0, n]$  – количество стимулов,  $k = [0, m]$  – количество испытуемых. Суммируя по  $j$ , получаем двумерную матрицу  $A_{ik} = \sum A_{ijk}$ ,  $k$ -я строка которой представляет собой набор оценок, данных  $k$ -м испытуемым стимулам.

В случае прямых методов оценки мы сразу получаем аналогичную матрицу непосредственных оценок.

Сравнивая распределения оценок по стимулам у разных испытуемых, мы можем разбить общую выборку испытуемых на подвыборки, в которых различия в распределениях оценок по стимулам незначительны в пределах выбранного уровня значимости.

Рассмотрим матрицу  $A'_{ik}$ , составленную из оценок, данных испытуемыми из выделенной однородной подвыборки.  $i$ -й столбец этой матрицы представляет собой случайную величину, каждая реализация которой – это оценка, данная  $i$ -му стимулу  $k$ -м испытуемым. Наша подвыборка сформирована из исходной выборки по принципу малых отличий между распределениями оценок по стимулам у испытуемых, т. е. малых различий между строками матрицы  $A'_{ik}$ , следовательно, можно ожидать, что случайные величины, представленные столбцами матрицы, будут иметь выраженный максимум. Таким образом, каждому стимулу можно сопоставить шкальное значение, найденное путем усреднения по однородной выборке испытуемых  $A'_i = \sum A'_{ik} / m$ , и оценить как погрешность оценки стимула отдельным испытуемым, так и погрешность среднего; т. е. использовать традиционный аппарат обработки экспериментальных данных.

В методе непосредственной количественной оценки испытуемый распределяет стимулы (дает оценки) по заданному количеству категорий, стараясь, как это требует инструкция, выдерживать интервалы между категориями психологически равными. Предполагается, что испытуемый способен выполнить это требование, это является основанием считать построенные шкалы шкалами интервалов.

В методе парных сравнений испытуемый распределяет между стимулами заданное количество баллов (количество выигрышей). При этом можно утверждать, что получаемая таким образом шкала является шкалой порядка.

Единый концептуальный подход к построению шкал на основе данных непосредственной количественной оценки и метода парных сравнений позволяет сравнить эти шкалы, определить вид допустимого преобразования, следовательно, тип полученных шкал.

Поставленная задача решалась, в частности, путем апробации и сравнения различных эмпирических процедур психологического измерения сложных перцептивных стимулов. Этим обусловлен выбор стимульного материала. С одной стороны, изменение каких-то параметров в стимульном наборе должно вызывать явные изменения реакций испытуемых, с другой стороны, стимулы должны провоцировать наиболее полное и разнообразное включение психологических функциональных систем в деятельность оценки.

Одним из самых значимых для человека элементов в системе его восприятия является человеческое лицо – центр приема и передачи социальных сигналов. Известно, что уже новорожденные предпочитают лицо или его схематическое изображение [7]. Исследования показали правомочность применения символических изображений лица в шкалах оценки выражения эмоций [6]. Ягер и Борг для построения подобных шкал использовали схематические изображения лица, в которых варьировалась степень изгиба дуги, представляющей на рисунках рот [5].

Более 90 % сюжетов любительской фотографии содержат изображение человека как самой значимой фигуры, определяющей смысл сюжета. В условиях тахистоскопического предъявления изображения человека воспринимаются при значительно меньшем времени предъявления, чем сюжеты без человека [2].

Исходя из вышеизложенного, художнику было дано задание сконструировать ряд изображений, в которых был бы осуществлен переход от схематического изображения предмета к стилизованному изображению лица через ряд изображений разной степени неопределенности. Полученные семь изображений предъявлялись испытуемым в качестве стимульного материала.

В психологическом измерении рисунков были использованы два критерия оценки. Первый критерий – интегральный, который выражает общее отношение испытуемого к данному изображению. В инструкции испытуемым предлагалось оценить свое общее отношение к стимулам, насколько они «предпочитают» эти стимулы другим. Поэтому в дальнейшем будем называть этот критерий критерием «предпочтение».

Второй критерий – частный. Испытуемых просили оценить в стимулах-рисунках степень похожести рисунка на живое человеческое лицо. Далее этот

критерий будем называть «похожесть». В пилотажных исследованиях в инструкции по оценке по второму критерию отсутствовало слово «живое». Испытуемых просто просили оценить степень похожести рисунка на лицо. Но испытуемые, легко разгадав замысел художника, давали очевидные одинаковые оценки. После того как в инструкцию ввели слово «живое», в оценках испытуемых появилось больше разнообразия. Например, для некоторых испытуемых первое изображение оказалось менее похожим на живое лицо, чем второе или даже третье.

Мы использовали простой несбалансированный метод парных сравнений (МПС) с принудительным выбором, т. е. испытуемый должен был в любом случае принять решение в пользу какого-нибудь стимула в паре, нейтральные ответы запрещались. Если испытуемый затруднялся, ему советовали делать выбор наугад, полагаясь на «внутренний голос».

Предъявлялись семь стимулов, следовательно, производился 21 акт сравнения. Для каждого испытуемого заполнялась матрица ответов обычным способом. Затем для каждого стимула суммировались его победы, и, таким образом, получали распределение «сырых» оценок между стимулами данным испытуемым.

Метод непосредственной количественной оценки (НКО). В этой эмпирической процедуре испытуемым предлагалось оценить степень выраженности критерия в стимулах по стобалльной шкале, причем концы шкалы не закреплялись за лучшим и, соответственно, худшим стимулами в наборе. Испытуемые получали инструкцию стараться использовать психологически равномерную субъективную шкалу, т. е. им указывалось, что психологическое расстояние, например, между 26 и 29 баллами равно расстоянию между 72 и 75 баллами. Выбор размаха шкалы от 0 до 100 объясняется желанием, во-первых, повысить точность измерения, во-вторых, дать испытуемому большую свободу действий.

Непосредственная количественная оценка субъективных интервалов (НОИ). Процедура НОИ была организована следующим образом. При проведении МПС испытуемому давалась дополнительная инструкция: «После того, как Вы в каждом акте сравнения выбрали тот стимул в паре, который, по вашему мнению, более соответствует предложенному критерию, оцените, на сколько баллов выражен критерий в этом стимуле больше по сравнению с другим стимулом в паре. Оценку сделайте в равномерной шкале размахом от 0 до 100 баллов.

Запишите оценку в соответствующую клеточку матрицы МПС. Если Вы считаете, что стимулы в паре никак не отличаются друг от друга, сделайте выбор стимула наугад, полагаясь на свой «внутренний голос», и в соответствующую клеточку матрицы впишите 0».

Процесс психологического измерения активизирует иерархический набор психологических систем, иерархия которого определяется, с одной стороны, объектом психологического измерения (стимульным набором и инструкцией), с другой стороны, индивидуальными особенностями испытуемого, т. е. непосредственно предметом психологического измерения, как мы его определили выше.

Разбиение выборки испытуемых на четыре подгруппы (см. ниже) демонстрирует четыре типа реакции на данный стимульный набор. Для того чтобы выяснить, имеют ли эти выделенные различия в типах реакции какое-либо пси-

хологическое основание или являются просто артефактом эксперимента, мы использовали ряд психометрических методик.

Влияние личностных особенностей на оценку исследовалось с помощью 16-факторного личностного опросника Кеттелла, опросника ПДТ Мельникова-Ямпольского и опросника Айзенка.

Влияние эмоционального состояния исследовалось с помощью шкалы самооценки Спилбергера – Ханина.

Участие мыслительно-познавательных процессов исследовали с помощью «прогрессивных матриц Равенна» и теста Д. В. Сочивко на определения индивидуального стиля познавательной деятельности (ИСПД) [1].

Особенности когнитивного стиля исследовались с помощью теста Струпа, теста скрытых фигур Терстона, теста свободной сортировки Гарднера, теста ригидности мышления Лачинса.

В эксперименте принимали участие в общей сложности более 1200 испытуемых – мужчин и женщин в возрасте от 16 до 45 лет, жителей разных городов страны, имеющих образование не ниже среднего. Все они приняли участие в основной фазе эксперимента, а именно – оценили стимульный набор по критерию «предпочтение» методами НКО (1010 испытуемых) и МПС (1147). По критерию «похожесть» оценили рисунки методом НКО 274 испытуемых, методом МПС – 324. Непосредственно субъективные интервалы (НОИ) оценили по критерию «предпочтение» 214 чел., по критерию «похожесть» – 93.

Из психометрических методик предлагались: ПДТ – 58 испытуемым, опросник Кеттелла – 40, опросник Спилбергера – Ханина – 125, ИСПД – 214, опросник Айзенка – 208, прогрессивные матрицы Равена – 63, методики определения когнитивного стиля – 128 испытуемым.

Количество испытуемых, привлеченных к участию в базовом эксперименте, определилось стремлением достичь компромисса между эффектами малой и большой выборки, стремлением довести объем самой малочисленной группы  $K_1$  (см. ниже) испытуемых до величины, позволяющей корректно применять методы статистического анализа (примерно 80 испытуемых).

Разное количество испытуемых, выполнявших разные психометрические методики, объясняется следующими соображениями.

Опросники Кеттелла и ПДТ применялись в пилотажном исследовании для определения дальнейшего направления работы, поэтому выборки объемом примерно 50 испытуемых вполне достаточно для проведения общего корреляционного анализа. Эти методики практически дублируют друг друга, поэтому они были использованы для взаимного контроля результатов.

Опросники Айзенка и Спилбергера – Ханина, а также методы исследования когнитивного стиля использовались для целенаправленного изучения действия соответствующих факторов в выделенных группах испытуемых.

Тест ИСПД был разработан относительно недавно, еще не получил всеобщего признания, поэтому к выполнению этой методики было привлечено такое большое количество испытуемых.

Объем выборки для установления возможных связей оценок стимулов с когнитивными стилями вполне достаточен для корреляционного анализа.

Исходным материалом для первичной обработки результатов эксперимента служили матрицы «сырых» оценок  $A_{ik}$ ,  $i = [1, n]$ ,  $k = [1, m]$ , где  $n$  – количество стимулов,  $m$  – количество испытуемых. Таким образом, строка матрицы – распределение «сырых» оценок, данных  $i$ -м испытуемым стимульному набору. Каждый элемент матрицы представляет оценку, данную испытуемым стимулу.

В результате применения эмпирических процедур получены наборы «сырых» оценок. На основе этих «сырых» оценок были построены шкалы в едином концептуальном и алгоритмическом ключе. Шкалы на основе оценок, полученных методами МПС и НКО, были построены способом, описанным выше.

Шкалы для оценок методом НОИ были построены способом, аналогичным построению шкал в модели Терстона. Только для этого использовалась матрица непосредственной оценки интервалов между стимулами. В этой матрице в пустые клетки вписывались соответствующие значения из «зеркально» расположенных клеток с обратным знаком. Таким образом, получалась полная матрица оценок интервалов, в диагоналях которой стояли 0. Далее суммировались элементы в столбцах матрицы, которые представляли шкальные значения стимулов. Операция усреднения по количеству стимулов и сдвига точки отсчета шкалы в 0 считалась излишней в соответствии со свойствами предполагаемой шкалы интервалов.

Чтобы свести к минимуму интериндивидуальные различия в шкалах, связанные с используемым методом, например с выбором испытуемым начала отсчета и размаха шкалы в методе НКО, и получить возможность более корректного сравнения шкал, полученных разными эмпирическими процедурами, индивидуальные шкалы были переведены в единую шкалу с помощью линейного преобразования:

$$y = (b_2 - b_1)(x - a_1) / (a_2 - a_1) + b_1,$$

где  $x$  – шкальное значение исходной шкалы;  $a_2$ ,  $a_1$  – минимальное и максимальное значения исходной шкалы соответственно;  $y$  – шкальное значение новой шкалы;  $b_2$ ,  $b_1$  – минимальное и максимальное значения новой шкалы, причем выбрано  $b_2 = 10$ ,  $b_1 = 0$ .

Шкальное значение для  $i$ -того стимула определяется как усредненная «сырая» оценка по всей выборке испытуемых:  $A_i = \Sigma A_{ik} / m \pm st / m$ , где  $s$  – оценка среднеквадратичного отклонения,  $m$  – объем выборки испытуемых,  $t$  – значение распределения Стьюдента, соответствующее выбранному уровню значимости. Эти данные приведены в табл. 1.

В оценках стимулов по критерию «похожесть» явно прослеживается тенденция к уменьшению оценок с ростом номера стимула, т. е. чем больше номер стимула, тем он меньше похож, по мнению испытуемых, на живое человеческое лицо.

Несмотря на некоторый разброс оценок, выборка испытуемых при оценках по критерию «похожесть» оказалась достаточно однородной и выделить какие-либо подгруппы не удалось.

Таблица 1

Шкальные значения, вычисленные по всей выборке испытуемых по данным МПС, НКО, НОИ для критериев «предпочтение» и «похожесть»

«Предпочтение»								«Похожесть»							
МПС															
<i>n</i>	1	2	3	4	5	6	7		1	2	3	4	5	6	7
<i>X</i>	6,1	5,5	3,7	6,1	3,3	4,7	5,6		8,9	7,7	5,7	5,5	4,6	2,0	0,7
$\sigma$	3,1	3,2	2,8	3,3	3,2	3,3	3,7		2,4	2,1	1,8	2,0	2,2	1,7	1,9
<i>m</i>	1123							324							
НКО															
<i>n</i>	1	2	3	4	5	6	7		1	2	3	4	5	6	7
<i>X</i>	6,0	5,2	3,3	5,3	2,9	4,3	5,2		8,9	7,6	5,6	5,6	4,5	1,4	0,5
$\sigma$	4,0	3,4	2,8	3,5	3,1	3,6	4,1		2,6	2,4	2,4	2,6	2,7	1,9	1,8
<i>m</i>	1010							274							
НОИ															
<i>n</i>	1	2	3	4	5	6	7		1	2	3	4	5	6	7
<i>X</i>	6,1	6,2	4,1	5,7	3,8	3,9	3,9		9,2	7,2	5,0	4,8	3,8	1,7	1,1
$\sigma$	4,1	3,1	2,8	3,2	2,9	3,5	3,8		2,4	2,0	2,4	1,9	2,0	2,2	2,5
<i>m</i>	214							93							

*X* – шкальное значение;  $\sigma$  – среднееквадратичное отклонение; *m* – объем выборки испытуемых, *n* – номера стимулов-рисунков.

Иная картина в случае критерия «предпочтение». В таблице 1 приведены шкальные значения для оценок по этому критерию. Распределение оценок по стимулам показывает, что с учетом погрешности измерения оценки стимулов мало отличаются. Можно прийти к выводу, что в среднем для испытуемых все стимулы почти одинаково предпочтительны (или, соответственно, неpreferи- тельны), но вид индивидуальных диаграмм распределения оценок говорит, что это не так.

Матрицу индивидуальных шкальных значений метода МПС подвергли факторному анализу методом главных компонент.

Были выделены два фактора, получившие разумное объяснение. Весовые значения для этих двух факторов (главных компонент) приведены в табл. 2.

Таблица 2

Значения весов для двух главных компонент, получивших интерпретацию

1. $D = 35,6$								2. $D = 18,2$							
	1	2	3	4	5	6	7		1	2	3	4	5	6	7
$W_1$	.71	.72	.53	-.12	-.14	-.77	-.76	$W_2$	-.45	-.23	.36	.63	-.60	-.15	-.41

$W_1$  – распределение по стимулам весов, определяющих фактор «похожести на человеческое лицо»;  $W_2$  – распределение весов по стимулам для фактора «неопределенность»;  $D$  – значение дисперсии.

Первый фактор можно интерпретировать как соответствующий степени похожести на лицо, так как значения весов практически повторяют порядок шкальных значений стимулов по критерию «похожесть» (см. табл. 1). Второй фактор можно интерпретировать как фактор неопределенности рисунка. В этом случае наибольший вклад в этот фактор дают «переходные» стимулы (3, 4), а наименьший – стимулы, на которых явно изображены или лицо, или предметы (1, 2, 6, 7).

В соответствии с выделенными факторами общая выборка испытуемых была разбита на две группы, которые получили названия «группа социально ориентированного типа», (гр.  $S$ ) и «группа когнитивно ориентированного типа», (гр.  $K$ ). Эти группы, в свою очередь, были разделены на две подгруппы каждая – гр.  $S_1$ , гр.  $S_2$ , гр.  $K_1$ , гр.  $K_2$  соответственно.

Разбиение осуществлялось следующим образом. Были вычислены суммы квадратов разностей рангов шкальных значений каждого испытуемого и рангов весов для каждого фактора для гр.  $S_1$  и гр.  $K_1$ , для гр.  $S_2$  и гр.  $K_2$  между рангами шкальных значений и рангами весов, взятых с обратным знаком. В соответствующие группы были отнесены испытуемые, у которых эти разности не превышали 26, что отвечает уровню значимости 5 %.

Для каждой выделенной группы были построены шкалы по данным используемых эмпирических методов (МПС, НКО, НОИ). Шкальные значения приведены в табл. 3.

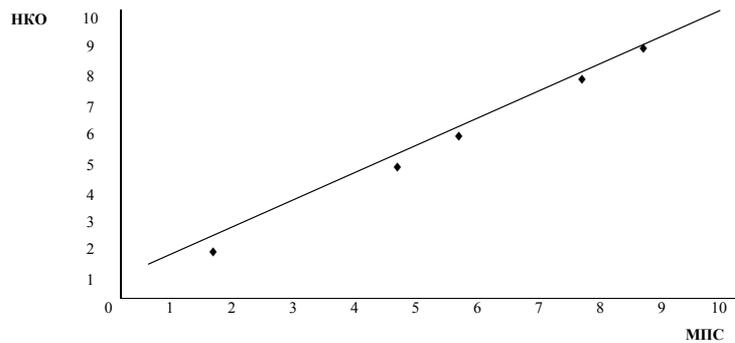
Таблица 3

Шкальные значения оценок по критерию «предпочтение» в выделенных группах испытуемых

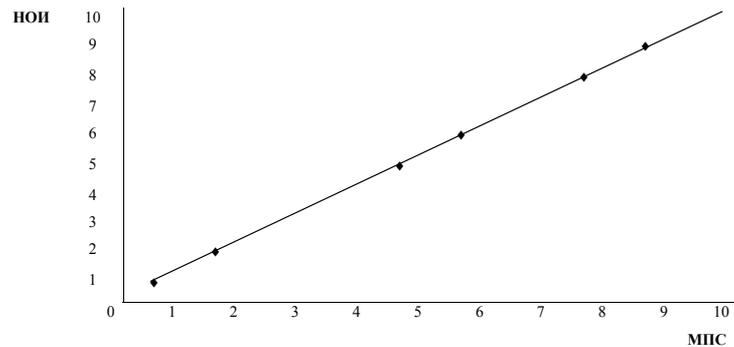
МПС								НКО							
$n$	1	2	3	4	5	6	7		1	2	3	4	5	6	7
$X$	9,1	8,2	5,5	5,4	3,3	1,5	1,7		8,7	7,4	4,7	4,8	2,8	1,5	1,8
$\Sigma$	1,6	1,4	2,2	2,7	2,5	1,5	2,2		2,3	2,4	2,4	3,0	2,7	2,7	2,9
$m$	315							Гр. $S_1$	282						
$X$	2,1	2,5	2,5	6,9	4,4	7,9	8,5		2,3	2,7	2,2	5,8	3,8	7,1	8,0
$\Sigma$	2,1	2,0	2,4	2,9	3,2	1,9	1,9		3,1	2,6	2,4	3,5	3,2	2,9	2,8
$m$	248							Гр. $S_2$	220						
$X$	1,9	4,5	5,7	9,3	7,3	4,2	2,5		3,1	4,2	4,8	8,0	6,0	4,2	3,9
$\Sigma$	2,3	3,3	2,6	1,3	2,1	3,0	2,5		3,6	3,3	3,0	2,9	3,4	3,3	3,6
$m$	84							Гр. $K_1$	77						
$X$	8,2	6,0	2,5	5,6	1,6	5,5	8,2		8,0	5,5	2,4	2,8	1,6	5,0	7,1
$\sigma$	2,0	2,7	2,3	2,2	2,1	2,9	2,1		2,0	3,0	2,3	2,8	2,2	3,3	3,5
$m$	234							Гр. $K_2$	208						
«Предпочтение» НОИ															
$n$	1	2	3	4	5	6	7		1	2	3	4	5	6	7
$X$	9,1	8,1	5,4	4,9	3,2	1,0	1,0		1,1	2,5	2,9	6,5	5,1	7,5	7,8
$\sigma$	2,0	1,7	2,0	2,5	2,2	1,2	1,8		2,1	2,5	2,5	3,0	3,2	2,1	2,8
$m$	89; Гр. $S_1$								42; Гр. $S_2$						
$X$	2,7	5,7	5,0	9,1	6,2	3,9	2,4		7,6	5,3	3,3	2,2	2,3	5,1	7,6
$\sigma$	3,3	3,6	2,6	1,8	2,7	3,3	2,8		2,7	3,0	2,9	2,3	3,0	3,1	2,8
$m$	16; Гр. $K_1$								23; Гр. $K_2$						

$X$  – шкальное значение;  $\sigma$  – среднееквадратичное отклонение;  $m$  – объем выборки (количество испытуемых в группе).

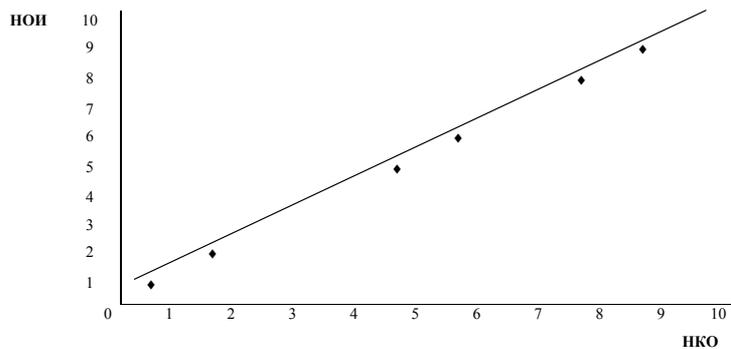
На рисунке 1 показаны примеры связи между шкалами, полученными МПС, НКО и НОИ для выделенных групп и разных критериев. Видно, что точки практически лежат на прямой линии. Шкалы, не представленные ради экономии места на рисунках, обнаруживают также линейную связь. Регрессионный анализ не дал повода отвергнуть гипотезу о линейном виде зависимости.



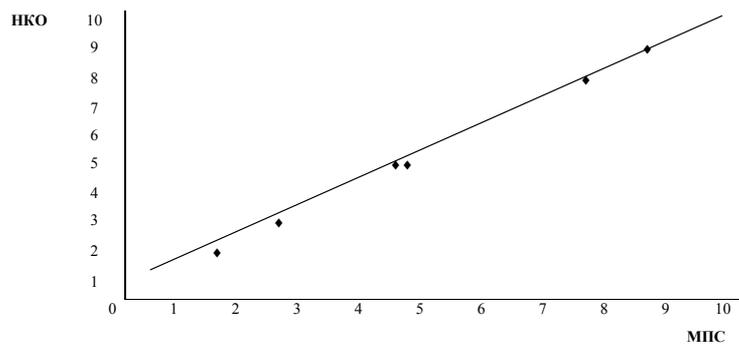
a – «Похожесть»



б – «Похожесть»



в – «Похожесть»



г – «Предпочтение», гр. S<sub>I</sub>

Рис. 1. Примеры связи между шкальными значениями, построенными по данным различных эмпирических процедур

В таблице 4 приведены надлежащие математические выкладки и значения оценок статистики Фишера.

Таблица 4

Оценка линейности зависимости между шкалами МПС и НКО в выделенных группах для критерия «предпочтение» и критерия «похожесть»

	Гр. $S_1$	Гр. $S_2$	Гр. $K_1$	Гр. $K_2$	«Похожесть»
$F_1$	0,06	0,07	0,06	0,09	0,09
$F_2$	121	59	24	59	121
$r$	0,98	0,96	0,91	0,96	0,98

$F_1, F_2$  – статистики Фишера для проверки адекватности линейной модели,  $F_3$  – статистика для проверки корректности линейной модели,  $r$  – коэффициент корреляции между шкальными значениями.

Таблица 5

Оценка линейности зависимости между шкалами МПС и НОИ в выделенных группах для критерия «предпочтение» и критерия «похожесть»

	Гр. $S_1$	Гр. $S_2$	Гр. $K_1$	Гр. $K_2$	«Похожесть»
$F_1$	0,29	0,13	0,1	0,05	0,09
$F_2$	59	46	68	80	121
$r$	0,96	0,95	0,94	0,97	0,98

$F_1, F_2$  – статистики Фишера для проверки адекватности линейной модели,  $F_3$  – статистика для проверки корректности линейной модели,  $r$  – коэффициента корреляции между шкальными значениями.

Таблица 6

Оценка линейности зависимости между шкалами НКО и НОИ в выделенных группах для критерия «предпочтение» и критерия «похожесть»

	Гр. $S_1$	Гр. $S_2$	Гр. $K_1$	Гр. $K_2$	«Похожесть»
$F_1$	0,22	0,18	0,14	0,22	0,23
$F_2$	80	32	24	16	46
$r$	0,397	0,93	0,91	0,87	0,95

$F_1, F_2$  – статистики Фишера для проверки адекватности линейной модели,  $F_3$  – статистика для проверки корректности линейной модели,  $r$  – коэффициент корреляции между шкальными значениями.

### Примечания к табл. 4, 5, 6

Проверяется линейная модель  $y = ax + b + e$ ;  $a, b$  – параметры модели,  $e$  – случайное отклонение.

#### 1. Проверка адекватности модели

Для этого была использована схема, в которой  $x$  и  $y$  не случайны, но могут быть измерены с некоторыми случайными ошибками, главное, чтобы ошибки были меньше диапазона изменений переменных. Исходная «балльная модель» удовлетворяет требованиям этой схемы:

$$F_1 = (m - 1)(1 - r^2) S_y^2 / [(m - 2) \sum S_{yi}^2]; F_2 = nF_1,$$

$m = 7$  (количество точек);  $S_{yi}$  – оценки соответствующих среднеквадратичных отклонений в точках  $i$ ;  $r$  – коэффициент корреляции между  $x$  и  $y$ ;  $n$  – количество измерений в точке  $i$ ;  $S_y$  – оценка среднеквадратичного отклонения переменной  $y$ .

Если  $F_1 > 1$ ,  $F_2 > F_0$ , гипотеза о линейной зависимости отвергается.  $F_0 = 4,2$  для уровня значимости 0,001.

## 2. Проверка значимости регрессии (корректности линейной модели)

$$F_2 = (m - 2) r^2 / (1 - r^2).$$

Если  $F_2 < F_0$ , гипотеза о корректности модели отвергается.  $F_0 = 6,6$  для уровня значимости 0,05.

Таким образом, нет оснований отказаться от гипотезы о линейной зависимости между шкалами, построенными по данным, полученным различными эмпирическими процедурами – МПС, НКО, НОИ. Следовательно, допустимым преобразованием для этих шкал является линейное, а именно линейное преобразование определяет шкалы интервалов.

Линейная связь между шкальными значениями показывает, что, с одной стороны, испытуемые действительно могут непосредственно оценивать субъективные расстояния между сложными перцептивными стимулами при работе с критериями разной степени сложности, с другой стороны, еще раз подтверждает, что построенные шкалы являются шкалами интервалов.

Считается общепризнанным, что МПС более точен, так как лежащая в его основе операция установления различия не вызывает у испытуемых таких затруднений, как операция установления равенства, на которую опирается метод НКО.

Очевидно, что процедура МПС очень громоздка, утомительна как для испытуемого, так и для экспериментатора, причем время проведения эксперимента резко возрастает с увеличением количества стимулов. Процедура НКО менее утомительна, время ее проведения минимально.

На рисунке 2 изображены автокорреляционные функции оценок для обоих критериев и обсуждаемых процедур.

Корреляции между оценками соседних стимулов в методе НКО значительно выше, чем в МПС, и практически все они (кроме одного случая) превышают линию 5%-ного уровня значимости. Вид автокорреляционных функций для простого и сложного критериев практически совпадает. Следовательно, на оценку каждого отдельного стимула влияет весь стимульный набор, и изменение стимульного набора (например, введение новых стимулов) может повлиять на оценку отдельного стимула. Как показали беседы с испытуемыми, последние пытались часто оценивать не стимул как таковой, а сначала разгадать закономерность построения ряда стимулов, воспринимали процесс оценки как игру, навязанную им экспериментатором.

Автокорреляционная функция МПС показывает, что оценки, полученные этим методом, не так сильно взаимосвязаны, как в методе НКО, но тем не менее говорить об их полной независимости нельзя. В одном из экспериментов испытуемым предлагались стимулы в процедуре МПС, но при этом, кроме двух стимулов сравнимой пары, весь остальной стимульный набор также находился в поле зрения испытуемого. Результаты этого эксперимента никак

не отличались от данных традиционного проведения МПС. Следовательно, в акте сравнения внимание испытуемого локализовано на предлагаемой паре, а остальные стимулы не замечаются. Но вид автокорреляционной функции все-таки указывает на влияние на оценку всего стимульного набора, возможно, через эталоны, формируемые в обобщенном образе испытуемого.

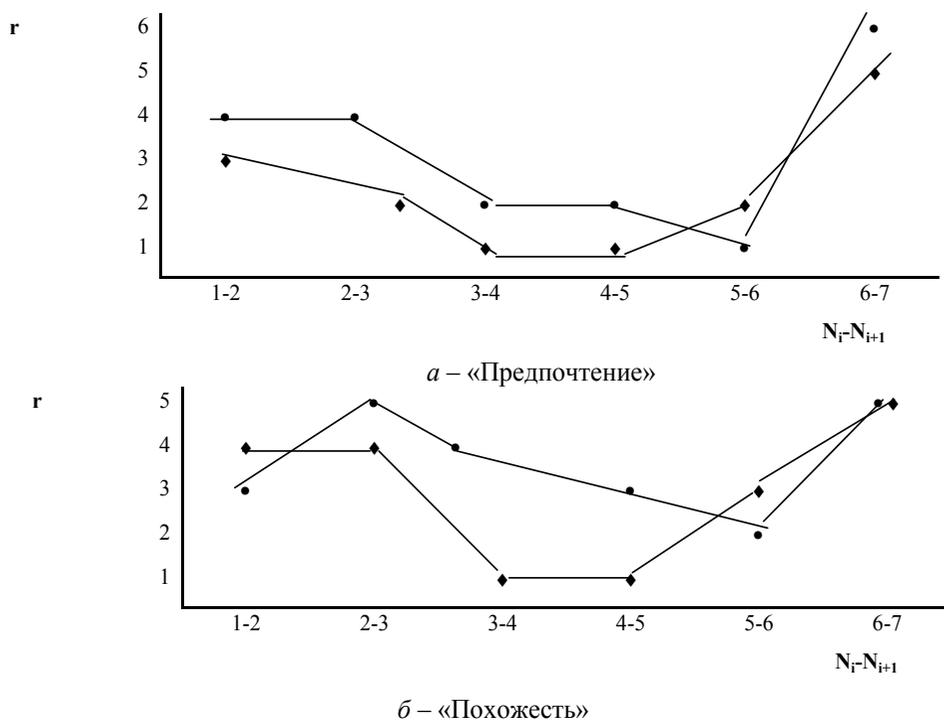


Рис. 2. Автокорреляционная функция для оценок методом НКО (•) и МПС (♦), пунктирная линия обозначает 5%-ный уровень значимости

Таким образом, выбор эмпирической процедуры зависит от конкретной экспериментальной ситуации, но для повышения валидности измерения полезно применять обе эмпирические процедуры, если это возможно, и строить по этим данным обобщенную шкалу.

До сих пор, рассуждая о типе построенных шкал, мы ни разу не упомянули о такой необходимой, по мнению многих, характеристике шкал интервалов, как единица измерения субъективного расстояния. Именно эта характеристика позволяет отнести шкалы интервалов к метрическим шкалам. Можно было бы выбрать в качестве единицы измерения расстояния ошибку оценки шкального значения, тем более что эту ошибку можно трактовать как оценку величины разностного порога (величины едва заметного различия – ЕЗР).

Это предлагает, например, А. А. Каралин и называет интервал неразличения субъективной единицей измерения (СЕИ). Он считает, что с ростом величины стимула возрастает величина используемых категорий [3].

При методе категорий равенство категорий задается инструкцией – это осознанный процесс, а рост категорий с ростом стимула – процесс неосознанный. Это же относится и СЕИ, и к ЕЗР. А. А. Каралин приходит к следующим выводам:

1) величина СЕИ изменяется пропорционально величине измеряемого стимула;

2) сходство категориальных шкал и шкал СЕИ показывает, что в их основе лежит динамика изменения одних и тех же психологических образований [7].

Но многие исследователи не считают возможным использовать оценку ошибки измерения (интервал неопределенности) в качестве единицы измерения [26].

Можно ли принимать за единицу измерения даже субъективную «мнимую» величину, значение которой не отражается на субъективную ось, человек ее «не видит»?

Наши данные показывают, что величина ошибки мало меняется от оценки к оценке по всей шкале. Таким образом, взяв среднюю по шкале ошибку, можно построить шкалу интервалов.

Но имеет ли смысл усложнять и без того очень громоздкий аппарат построения психологической шкалы? Рассмотрим вид допустимого преобразования, определяющий тип шкалы интервалов:  $y = ax + b$ . По определению шкалы интервалов задается с точностью до начала отсчета – параметр  $b$  и с точностью до масштаба – параметр  $a$ .

Последнее положение означает, что выбор единицы измерения меняет шкалу, но чтобы новая шкала оставалась шкалой интервалов, необходимо сохранение жесткой структуры шкалы, т. е. чтобы все расстояния (интервалы) в новой шкале изменились в одинаковое число раз. Из этого следует, что если мы имеем два разных стимульных набора, соответственно, две шкалы интервалов, каждую со своей структурой, то, чтобы построить единую шкалу, объединяющую оба набора стимулов, необходимо, чтобы в одном наборе имелось как минимум два стимула из другого набора. Это позволит нам определить взаимный масштаб, коэффициент перевода шкальных значений одной шкалы в другую. Таким способом нам удалось построить шкалу интервалов оценки качества более ста любительских фотоотпечатков, измеренных методом парных сравнений [6].

Итак, выбор единицы измерения весьма произволен при соблюдении условий, не разрушающих структуру шкалы. Поэтому все шкальные значения в нашей работе приведены в условных единицах, тем более что любые манипуляции с оценками, имеющими собственные погрешности, сильно увеличивают окончательную погрешность измерения.

Чем обширнее класс высокоспецифических особенностей исследуемых явлений, тем более универсальными должны быть признаки, общие для них всех. Психические явления не составляют тут исключения. Иначе говоря, только самые универсальные характеристики психики – общие для всех многосторонних и многоаспектных частных и специфических ее проявлений, воплощенных в личности. Закон обратной пропорциональности объема и содержания: чем более частные и высокоспецифические характеристики должны быть

охвачены соответствующим методом измерения, анализа и психодиагностического заключения, тем более общий характер должна носить соответствующая система единиц измерения.

Аналогично системе единиц измерения в физике (сантиметр, грамм, секунда и т.д.) в основании системы психологических средств измерения, а затем и психологических единиц измерения должны лежать единицы, относящиеся к самым универсальным параметрам психики. Таким образом, родовая специфичность психической ткани может и должна быть выражена в единицах измерения особенностей структуры психического времени, психического пространства, специфических форм выражения психологической интенсивности (психологической энергии) и психологической качественной специфичности.

Рассмотрим трехмерную матрицу  $A_{ijk}$ , полученную в результате проведения процедуры МПС. Вычислим две величины:  $A_i = \sum \sum A_{ijk} / n$  и  $P_i = \sum \sum A_{ijk} / mn$ , где  $A_i$  – шкальное значение «балльной» модели,  $P_i$  – оценка вероятности предпочтения, средняя для данного стимула. Если разделить  $A_i$  на  $m$ , что, очевидно, не меняет структуру шкалы, то эти две величины численно совпадут.

Согласно закону сравнительных суждений Терстона, для построения шкалы интервалов вероятности предпочтения переводят в квантили нормального распределения.

В работе [6] приведен рисунок, который демонстрирует связь между шкалами, построенными на основе данных МПС, классическим методом Терстона и данных метода категорий в модификации, приближающей его к методу НКО. Точки на рисунке очень хорошо ложатся на кривую, напоминающую огиву (вероятностную кривую нормального распределения). Делается вывод, что шкалы, построенные по данным различных эмпирических процедур, обнаруживают линейной связи.

Вернемся к нашим результатам. В «балльной» модели шкальное значение шкалы НКО, с точностью до линейного преобразования, совпадает со шкальным значением МПС. Последнее, в свою очередь, численно совпадает с вероятностью предпочтения. Следовательно, шкальное значение НКО, с точностью до линейного преобразования, совпадает со значением вероятности предпочтения. Поэтому, переводя вероятности в квантили, мы получаем фактически связь между вероятностями и квантилями, т. е. кривую, тождественную огиве. Рисунок демонстрирует не различие между эмпирическими процедурами МПС и НКО, а различие между вычислительными процедурами – пример того, как выбор разных математических моделей искажает эмпирически полученную информацию.

Все наши выводы о типе построенных шкал опираются на эмпирически полученные данные, следовательно, их надежность ограничена погрешностью эксперимента. Вполне возможно, что уменьшение погрешности измерений может привести к отказу от модели линейной связи, которая лежит в основе наших рассуждений. Но погрешность измерений в наших экспериментах с методом НКО вполне сравнима с единицей измерения предложенной стобалльной шкалы. Увеличение точности возможно только за счет более подробной шка-

лы, а это не имеет смысла, так как литературные данные и наш собственный опыт говорят, что стобалльная шкала – почти предел чувствительности в таком рода измерениях.

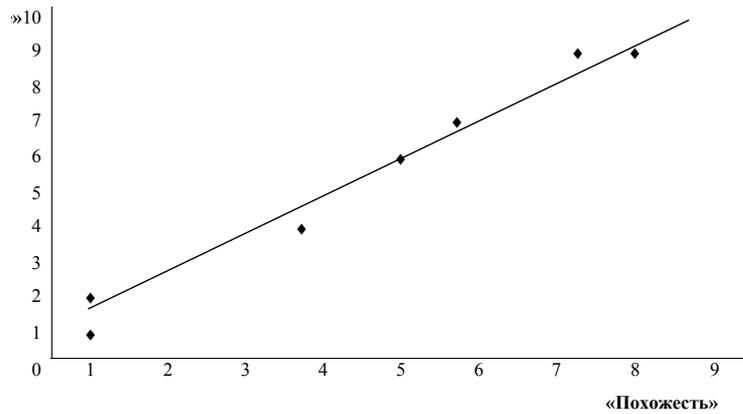
К. К. Платонов говорил, что самой сильной измерительной шкалой является трехбалльная (слабый, сильный, очень сильный; меньше, равно, больше). пятибалльную шкалу применять можно, а более многобалльная требует дополнительных «формулировок критериев для надежного применения». Он предлагал метод полярных баллов, который заключался в наложении двух полярных трехбалльных шкал, в результате чего получается пятибалльная шкала, не потерявшая силы трехбалльной.

Нами ранее было показано, что испытуемые не могут выдерживать пятибалльную шкалу психологически равномерной. Например, субъективное расстояние между 2 и 3 баллами оценивается значительно большим, чем между 3 и 4 баллами. Кроме того, стобалльный размах шкалы позволяет снизить влияние центростремительного эффекта, при котором испытуемые избегают давать стимулам крайние оценки, и предоставляет испытуемым достаточную свободу выбора. В описанном выше эксперименте 62 % испытуемых использовали «некруглые» оценки по критерию «предпочтения» и 54 % – по критерию «похожесть». То есть более сложный критерий требует более дифференцированной шкалы.

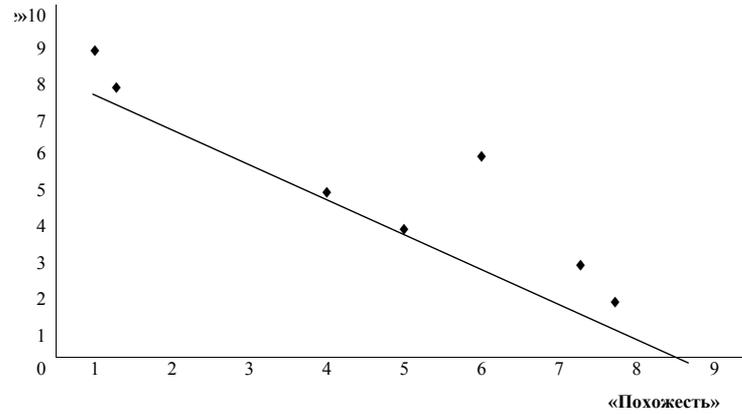
Установление типа шкал предоставило возможность построить графики функциональных зависимостей (одно из преимуществ шкал интервалов). На рис. 3 приведены графики зависимости степени выраженности интегрального критерия «предпочтение» от частного критерия «похожесть» в четырех выделенных группах. Приведены зависимости между шкалами, построенными по данным МПС. Для результатов применения других методов зависимости будут аналогичны из-за линейной связи между ними.

Психофизика давно научилась строить графики зависимости «субъективного» от «объективного», когда по осям отложены психофизические корреляты. Психофизические законы – логарифмический Фехнера и степенной Стивенса – представляют такие зависимости. На рисунке 3 приведены зависимости «субъективного» от «объективного».

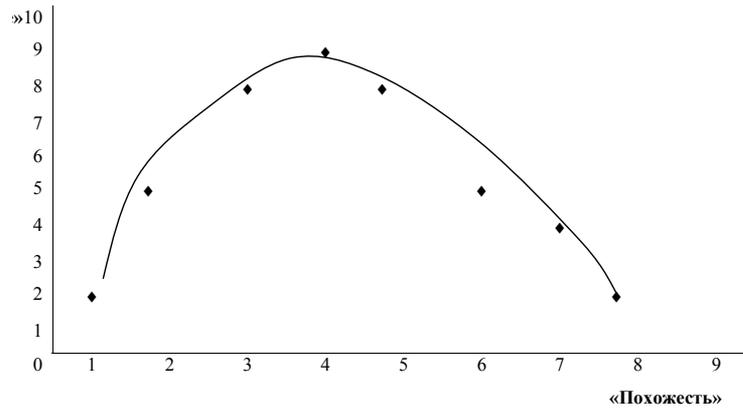
Проанализируем вид зависимостей. Графики первой и второй групп ( $S_1$  и  $S_2$ ) демонстрируют монотонную зависимость «предпочтения» от «похожести». Чем более стимул похож на лицо, тем он более предпочтителен в первой группе испытуемых и тем менее предпочтителен во второй. Можно предположить, что фактором, определяющим реакцию испытуемых в первой и второй группах на предложенный стимульный набор, было человеческое лицо, которое притягивает испытуемых первой группы и отталкивает испытуемых второй. Мы назвали этот тип реакции социально ориентированным типом.



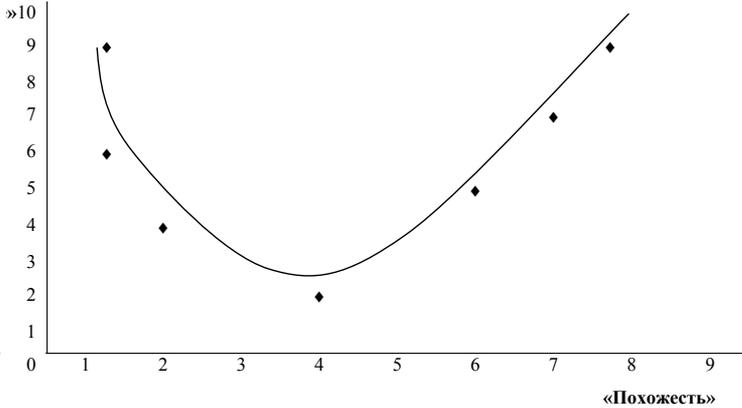
$a - \text{гр. } S_1$



$б - \text{гр. } S_2$



$в - K_1$



$г - \text{гр. } K_2$

Рис. 3. Зависимость оценок критерия «предпочтение» (ось ординат) от критерия «похожесть» (ось абсцисс) в выделенных группах

Графики третьей и четвертой групп ( $K_1$  и  $K_2$ ) имеют явно выраженный экстремум. В третьей группе испытуемые предпочитают неопределенные стимулы и отвергают равным образом явно похожие и непохожие на лицо. В четвертой группе, наоборот, испытуемые предпочитают или очень похожие на лицо стимулы, или совершенно непохожие, промежуточные изображения отвергаются (минимум на графике). Реакция испытуемых определяется не предметным содержанием стимулов-изображений, а степенью неопределенности стимула в стимульном наборе. Этот тип реакции мы назвали когнитивно ориентированным типом.

Каждый из двух выделенных типов содержит полярно противоположные тенденции: похожесть – непохожесть, определенность – неопределенность. В основе этих двух типов реакции лежат, по-видимому, разные психические процессы. Предположительно, эмоциональные процессы доминируют в реакциях первого типа, а познавательные – второго.

### Список литературы

1. Сочивко Д. В. Математические модели в психолого-педагогических исследованиях / Д. В. Сочивко, В. А. Якунин. – Л. : ЛГУ, 1988. – 68 с.
2. Худяков А. И. Психофизика обобщенного образа / А. И. Худяков // *Palmarium Academic Publishing*. LAP LAMBERT Academic Publishing GmbH & Co. KG. Saarbrücken. – Germany, 2012. – 247 p.
3. Худяков А. И. Психология измерений / А. И. Худяков ; РГПУ им. А. И. Герцена. – СПб. : Копи-Р Групп, 2013. – 220 с.
4. Худяков А. И. Экспериментальная психология / А. И. Худяков. – Харьков : Гуманитар. центр, 2016. – 408 с.
5. Borg G. Are we subjected to a «long-standing measurement oversight»? / Borg G. // *Fechner Day – 2001. Proceedings of the Seventeenth Annual Meeting of the International Society for Psychophysics* / ed. by E. Sommerfeld. – Leipzig, 2001. – P. 304–307.
6. Ekman P. Constants across cultures in the face and emotions / P. Ekman, W. V. Friesen // *Journal of Personality and Social Psychology*. – 1971. – N 17. – P. 124–129.
7. Sydow H. *Mathematische Psychologie* / H. Sydow, P. Petzold. – Berlin : Springer, 1981. – 323 p.

## Possibility of Constructing Interval Scale in Psychological Measurements

A. I. Khudiakov

*The Herzen State Pedagogical University of Russia, Saint-Petersburg*

**Abstract.** Basic problem of psychological measurements is disparity between psychological analysis of measurement process and corresponding mathematical models. Mathematical model of any process is valid if there is prior idea of the process nature. When using probability distributed with regard to the general principle there is no solution of the problem, the most problematic aspects are just avoided. In this case the subject of the research is handled as a black box. Using the results of direct methods of stimulus psychological measurements having no physical measures it is possible to construct a metric scale of intervals and rela-

tions. In the scales of this kind there is no particular unit of measurement since corresponding acceptable transformations enable to use a conditional unit.

**Keywords:** psychological measurement, psychophysical scaling, interval scales, empirical procedures, probability models, paired comparison method, direct qualitative estimation, direct qualitative estimation of subjective intervals.

*Худяков Андрей Иванович  
доктор психологических наук, профессор  
Российский государственный  
педагогический университет  
им. А. И. Герцена  
191186, г. Санкт-Петербург, 48,  
Набережная Реки Мойки, д. 48  
e-mail: haipsy@yandex.ru*

*Khudyakov Andrey Ivanovich  
Doctor of Sciences (Psychology), Professor  
Herzen State Pedagogical University of Russia  
48, Moika River Embankment,  
Saint-Petersburg, 191186  
e-mail: haipsy@yandex.ru*