

Психодиагностика

УДК 159.9.07 + 159.98
ББК Ю980.3

ИПСАТИВНЫЕ ЛИЧНОСТНЫЕ ТЕСТЫ ВЫНУЖДЕННОГО ВЫБОРА: СПЕЦИФИКА РАЗРАБОТКИ

А.Е. Иванов

Описаны подходы к изучению возможностей применения ипсативных данных в сфере психодиагностики, в частности, к возможности разработки тестов формата вынужденного выбора для нужд профотбора. Показано, что при создании такого рода методик необходимо учитывать некоторые специфические факторы, не оказывающие влияния на процесс разработки нормативных тестов. Приводятся результаты обзора зарубежных публикаций по созданию и использованию ипсативных тестов с пунктами в формате вынужденного выбора. Рассмотрены основные отличия и особенности разработки ипсативных тестов в сравнении с нормативными, описана специфика как ипсативных данных, так и ипсативных тестов вынужденного выбора. В качестве теоретической основы для сравнения процессов создания ипсативных и нормативных методик взята методология и технология последовательности разработки теста, описанная Н.А. Батуриным и Н.Н. Мельниковой в серии публикаций под названием «Технология разработки тестов».

Ключевые слова: ипсативный тест, ипсативные данные, ипсативные шкалы, формат вынужденного выбора, разработка психодиагностического теста.

В последнее десятилетие в англоязычной литературе по психодиагностике увеличивается число публикаций, описывающих поиск метода диагностики некогнитивных конструкторов, альтернативного по отношению к наиболее распространенным самооценочным нормативным опросникам в формате единичных стимулов (*single stimuli*) (Brown, Bartram 2009). Такого рода опросниковые методики, несмотря на свою популярность и проработанность теоретической основы, обладают недостатками. В первую очередь к ним относятся простота как умышленной, так и неосознанной фальсификации результатов таких методик со стороны испытуемого. В английской профессиональной терминологии этот феномен обозначается словом «*faking*» (поддельвание), а в русской чаще всего речь идет об «обмане» теста. Одним из способов контроля этого феномена является использование альтернативного класса методик, основанного на т.н. процедуре вынужденного выбора.

Отмечается разная сопротивляемость методик вынужденного выбора «обману» со стороны испытуемого, имеющая тому разные причины (Karpatschhof, Elkjaer, 2000; Martin, Bowen, 2002; Meade, 2004). Но результаты всех таких исследований сходятся на том, что по

сопротивляемости «обману» со стороны испытуемого методики с пунктами в формате вынужденного выбора демонстрируют лучшие результаты, нежели методики с пунктами в виде единичных стимулов. Однако у методик вынужденного выбора есть несколько отличительных черт, которые определяют область их применения, а также специфику их разработки и использования. Например, для ипсативного оценивания подходят только те области измерений, в которых респонденты способны осмысленно выбирать между альтернативами или ранжировать их в порядке предпочтения. Следовательно, для ипсативного оценивания не подходят области измерения, не связанные с системой предпочтений испытуемого. Но при этом ипсативный формат подходит для оценки мотивационной направленности, личностных предпочтений, ценностей и других конструкторов, в которых соотношение измеряемых характеристик имеет не меньшее значение, чем их абсолютная выраженность (Cornwell, Dunlap, 1994; Deaton, Glasnapp, 1980).

Для психологического измерения важными являются два условия: во-первых, все измерения феномена должны находиться в одном континууме, обладающем постоянной направленностью; во-вторых, такой контину-

ум должен быть одинаков для всех испытуемых (Cattell, 1944). Ипсативные тесты при конструировании требуют большего объема работы по проверке психометрических параметров отдельных ответов, пунктов и шкал, так как возникает задача избегания систематичности в операциональной (или внутренней) зависимости между шкалами (internal interdependence), неизбежно возникающей при объединении нескольких шкал в одном пункте (Brown, Bartram, 2009). Игнорирование этого требования может привести не только к сведению в один пункт неравноценных по социальной желательности или «трудности» стимулов, но и к нарушению первого из описанных ранее условий психологического измерения. В таком случае сравнение, являющееся главной операцией испытуемого при выполнении теста, окажется принципиально невозможным (Cattell, 1944; Karpatschhof & Elkjaer, 2000). В связи с этим важно указать на особые условия, которые необходимо учитывать на каждом из этапов разработки ипсативных методик, а также на отличия в разработке ипсативных и нормативных тестов.

Организационный и содержательный этап разработки методики

За исходную точку в описании особенностей разработки ипсативных опросников формата вынужденного выбора была взята система разработки нормативных психодиагностических методик, приведенная в серии статей Н.А. Батурина и Н.Н. Мельниковой «Технология разработки тестов» (Батурин, Мельникова 2009, 2010, 2011).

По этой технологии, одной из задач первых этапов разработки любой психодиагностической методики является обеспечение ее валидности, формулирование предварительных характеристик теста, определение исходного концепта. Ипсативные решения предъявляют особые требования к изучаемому конструкту, и, как следствие, к конструктивной валидности. В случае с ипсативными тестами конструктивная валидность в значительной мере зависит от связей между характеристиками или частями конструкта, измеряемого тестом.

Даже оптимальная согласованность и дискриминантность шкал теста не могут исправить неправомерно принятых допущений, составляющих базис измерительной процедуры. Ипсативное шкалирование предъявляет повышенные требования к операциональной валидности, так как априори требует приня-

тия большого количества базовых допущений, способных оказаться нерелевантными концепту измеряемой характеристики.

Например, ипсативный формат не подходит для методик, измеряющих единственную характеристику или нераздельный конструкт, поскольку ипсативные данные отражают соотношения, а не абсолютные величины. При этом возможно применение формата вынужденного выбора в одномерном измерении, особенно в том случае, когда с помощью методологии ИРТ в пункты объединяются высказывания со значительно различающейся «силой». В этом случае конечное соотношение выбранных относительно «сильных» и «слабых» высказываний может служить некоей мерой выраженности измеряемой латентной черты (Ackerman, 2005; Brown, Bartram, 2009; McCloy, Heggstad, 2005; Stark, Chernyshenko, 2005). Получаемые с рамках такого подхода данные уже не будут являться чисто ипсативными, но они все так же будут основываться на ипсативной информации, т. е. на субъективной позиции испытуемого.

Ипсативный формат не подходит для тех видов психодиагностического измерения, в которых важны абсолютные характеристики измеряемого конструкта или ранг испытуемого в генеральной совокупности иерархий выраженности оцениваемой латентной характеристики. Другими словами, все формы измерения с наличием правильных / неправильных вариантов выбора, а также все другие формы измерения с использованием внешнего объективного критерия не подходят для ипсативного инструментария. Главной особенностью ипсативных данных является то, что точкой отсчета в них является не внешний критерий, а субъективная система отсчета и оценки каждого испытуемого. Поэтому для обеспечения валидности такого инструмента необходимо точно определить, в какой мере субъективные представления испытуемого могут служить точкой отсчета для измеряемого конструкта (Bartram, 1996).

Другая важная особенность ипсативного инструментария заключается в том, что результаты отдельных шкал могут быть интерпретированы исключительно в рамках константного целого. В случае ипсативной методики нельзя ограничиться изучением только какой-то одной части конструкта – необходимо измерить характеристики всех составляющих его частей. Только такой подход может выявить истинные соотношения между час-

тями конструкта (Hough, 1998). Например, имеется некое целое, состоящее из 3 частей, находящихся в соотношении 1/2/7. Соответствующие доли этих частей будут составлять 0,1/0,2/0,7. Предположим, что истинная картина соотношения трех характеристик некоего конструкта именно такова. Если же произвести измерение соотношения частей того же конструкта, исключив третий, самый большой элемент, то получится отношение 0,33/0,66 – то есть отношение между оставшимися частями осталось прежним, но их значения по отношению к родовому конструкту значительно изменились, и внутренняя структура конструкта исказилась. Все показатели ипсативных тестов интерпретируются корректно только в рамках объединяющего их общего конструкта, и при исключении из измерения какой-либо его части неизбежно возникнут искажения в общей картине (Cattell, 1944).

С вышеприведенным условием связано то, что только немногие психологические концепты могут быть адекватно описаны математической моделью, предлагаемой ипсативным форматом. Подходящими при этом являются только концепты, указывающие на то, что конструкт состоит как минимум из двух независимых составляющих. При решении разработчиком методики задачи произвести с ее помощью межиндивидуальные сравнения необходимо буквально трактовать константность суммы частей изучаемого конструкта и негативные корреляции между ними. Это положение привносит особые требования к теоретическому оформлению конструкта. Например, негативные корреляции между некоторыми чертами личности иногда выявляются в эмпирических исследованиях, но нельзя сказать, что все черты личности негативно коррелируют друг с другом или, в крайнем случае, независимы. Из этих условий следует вывод, что интерпретация ипсативных показателей, необходимая для межиндивидуального сравнения испытуемых, предъявляет почти недостижимые требования к структуре изучаемого конструкта (Cornwell, Dunlap, 1994). В случае заинтересованности исследователя в выявлении соотношения компонентов изучаемого конструкта ипсативный формат является лучшим выбором. Естественно, что межиндивидуальное сравнение и стандартизация в таком случае будут валидными только при анализе общего профиля, а не отдельных шкальных оценок.

Выбор ипсативного формата требует от

исследователя четкого осознания не только целей и областей применения своей методики, целевой группы испытуемых, но и особенностей измеряемого конструкта и самого ипсативного формата. Разработка равных по психометрическому качеству нормативного и ипсативного психодиагностического инструментов значительно различается по затрате ресурсов. Создание ипсативной методики, сопоставимой по качеству со своим нормативным аналогом, является более сложным процессом, чем создание нормативной методики. Считается, что при решении задачи исключения ошибок измерения, связанных со стилем ответов испытуемого, следует использовать ипсатизацию (Cornwell, Dunlap, 1994).

Разработка и компоновка ипсативных пунктов

Разработка ипсативных методик на шаге формирования репрезентативной выборки содержания методик психологической диагностики личностной сферы уже само по себе является сложной задачей, однако в случае использования ипсативного формата при решении такой задачи добавляются новые проблемы, требующие отдельного внимания.

На этапе компоновки ипсативной методики особое внимание необходимо уделять контролю возникающей внутренней зависимости. В рамках ипсативного теста в формате вынужденного выбора необходимо достичь оптимального сочетания шкал в пунктах, т. е. высказывания или иные стимулы каждой шкалы должны равное число раз объединяться в пунктах с высказываниями или стимулами из других шкал методики. Это необходимо для выравнивания уровня так называемой «внутренней зависимости между шкалами на уровне пунктов» (Deaton, Glasnapp, 1980). Каждую шкалу теста должно составлять равное число высказываний. Кроме того, в рамках одного пункта все высказывания должны принадлежать одной содержательной области. Однако это правило часто игнорируется даже известными разработчиками ипсативных тестов (Meade, 2004). Если все высказывания или стимулы пункта относятся к одной содержательной области, то у испытуемого не будет возникать затруднений и дискомфорта при сравнении высказываний из качественно разных содержательных областей. В качестве примера можно привести два пункта гипотетической ипсативной методики, представленных в формате вынужденного выбора:

Пункт 1

Я предпочитаю активные формы отдыха.

Я предпочитаю в свободное время просто расслабиться.

Пункт 2

Я предпочитаю активные формы отдыха.

Меня часто посещает страх смерти.

Очевидно, что в пункте № 1 два высказывания относятся к одной предметной области и коррелируют отрицательно. В пункте № 2 высказывания относятся к разным предметным областям и поэтому теоретически могут коррелировать положительно. Выбор ответа в пункте № 2 не только может вызвать у испытуемого дискомфорт и потерю мотивации к тестированию, но и имеет меньшую ценность с точки зрения интерпретации результатов. Как следствие, при составлении таблицы спецификации соотношения содержательных элементов в тесте необходимо стремиться объединять в ипсативных пунктах стимулы не только равные по социальной желательности и «силе», но и относящиеся к одной предметной области (Hermelin, Robertson, 2001).

Следует отметить, что приведенное выше требование трудноосуществимо в отношении многих областей личностной диагностики, однако это не указывает на неверность самого правила, а ставит под вопрос использование ипсативного формата для диагностики таких конструктов (Hermelin, Robertson, 2001). Показанием к использованию ипсативного формата является также специфическая «сегментарная» структура диагностируемого конструкта.

На этапе разработки пунктов ипсативной методики представляется наименее трудозатратным использовать стимулы с уже проверенными психометрическими характеристиками. Но если требуется ввести новые стимулы, то, прежде чем компоновать их в пункты вынужденного выбора, следует проверить их характеристики в рамках пилотажного исследования. Такая проверка вызвана двумя факторами. Во-первых, объединение стимулов в пункты изменит их ранее наблюдавшиеся психометрические свойства, поскольку они станут частью большей психометрической единицы со своими внутренними взаимосвязями, требующими отдельного анализа. Во-вторых, корректное группирование высказываний требует знания некоторых характеристик (например, факторной нагрузки на разные шкалы методики, показателей надежности-согласованности, дискриминативности, уровня социальной желательности, «силы» или

порога ответа и других их характеристик в рамках IRT). Поскольку ипсативные данные ограничены в применимых к ним статистических методах анализа, характеристики всех стимулов и шкал должны быть определены заранее в рамках нормативной модели. Таким образом, разработка ипсативной методики частично включает в себя разработку аналогичной ей нормативной методики. Однако уже на этом этапе необходимо отбирать высказывания в соответствии с предполагаемой структурой итогового варианта методики. Все шкалы должны быть примерно равны по психометрическим характеристикам и содержать равное количество высказываний (Bartram, 2007).

Важным моментом также является определение числа стимулов, объединяемых в одном пункте. Если инструмент включает в себя от двух до четырех шкал, то существует возможность объединять в каждом пункте все шкалы методики. Но в то же время малое число шкал увеличивает долю внутренней зависимости на уровне ковариационной матрицы результатов методики (Hicks, 1970).

Специалисты британской компании SHL Group, занимавшиеся созданием ипсативных аналогов нормативных методик, пришли к выводу о том, что включение в ипсативную методику большого числа шкал (вплоть до нескольких десятков, как в случае с широко известной методикой OPQ32i) нивелирует многие из ограничений, существующих для результирующих ипсативных тестовых данных (Brown, Bartram, 2009). Действительно, использование большого числа шкал уменьшает внутреннюю зависимость на уровне ковариационной матрицы. Однако следует учитывать, что различия между ипсативными и нормативными данными имеют не только математическую природу (Cattell, 1944; Clemans, 1966; Hicks, 1970; Meade, 2004). Кроме того, увеличение числа шкал будет увеличивать не только сложность композиции теста, но и длительность его выполнения.

Известно, что процесс принятия испытуемым решения о том, какое из высказываний выбрать, не согласуется с одной из известных формул классической теории тестов:

$$X_j = T_j + e, \quad (1)$$

где X_j – полученная при помощи нормативной психометрической методики мера j -й черты испытуемого; T_j – истинная выражен-

ность латентной j -й черты; e – суммарная ошибка измерения.

В ипсативной методике на выбор испытуемого влияет определенное число факторов, при минимальном числе которых описывается формулой (Meade, 2004):

$$X_j = \frac{c + T_j + e_j - \sum_l^{s \neq j} (T + e)}{2}, \quad (2)$$

где X_j – полученная при помощи нормативной психометрической методики мера j -й черты испытуемого; c – число пунктов методики и число стимулов, направленных на измерение каждой части конструкта; T_j – истинные баллы измеряемого конструкта; e – ошибка измерения соответствующей шкалы;

$\sum_l^{s \neq j} (T + e)$ – сумма истинных баллов и ошибок измерения всех остальных частей измеряемого конструкта.

При достаточно большом числе факторов, влияющих на выбор испытуемого, формула имеет следующий вид (Meade, 2004):

$$X_{ijn} = (|T_j, T_k, \lambda_{i1j}, \lambda_{i2k}, e_{i1}, e_{i2}, \tau_n), \text{ если} \quad (3)$$
$$(T_j - \lambda_{i1j}) + e_{i1} + \tau_n > (T_k - \lambda_{i2k}) + e_{i2} + \tau_n,$$

где X_{ijn} – наблюдаемая реакция испытуемого n на 1-й стимул из приведенных в пункте i для измерения черты j ; T_j – истинное значение выраженности латентной черты j у испытуемого n ; T_k – истинное значение выраженности латентной черты k у испытуемого n ; λ_{i1j} – «сила» 1-го стимула или значение выраженности латентной черты j , необходимое для выбора стимула 1 из пункта i ; λ_{i2k} – «сила» 2-го стимула или значение выраженности латентной черты k , необходимое для выбора стимула 2 из пункта i ; e_{i1} – случайная ошибка измерения, присущая стимулу 1 пункта i ; e_{i2} – случайная ошибка измерения, присущая стимулу 2 пункта i ; τ_n – характеристики процесса принятия решения, присущие испытуемому n .

Таким образом, чем больше стимулов объединяется в ипсативном пункте, тем больше влияющих на выбор испытуемого факторов, тем сильнее проявляется внутренняя зависимость на уровне пунктов и тем сложнее процесс выбора со стороны испытуемого.

Из формул (2) и (3) следует, что в идеальном случае ипсативная методика должна включать в себя как можно больше шкал (для минимизации внутренней зависимости на уровне шкал), объединенных общим конст-

руктом, но в каждом пункте должно быть объединено минимально возможное число шкал.

Существует мнение, что влияние увеличения числа группируемых в одном пункте шкал не окажет существенного влияния на уровень внутренней зависимости между шкалами в рамках одного пункта. При этом сложность когнитивного процесса выбора будет возрастать, что, несомненно, увеличит долю ошибки измерения. Как следствие, выполнение ипсативной методики требует большого умственного напряжения со стороны испытуемого, чем выполнение аналогичной нормативной методики (Vasilopoulos, Cucina, Duomina, Morewitz, Reilly, 2006). При выполнении нормативных методик испытуемый должен перенести свою внутреннюю оценку стимула с субъективной шкалы на шкалу ответов методики (другими словами, он производит только одну операцию оценки). Если же испытуемому предлагается выбрать один из двух стимулов, то ему требуется найти подходящую субъективную шкалу, объединяющую эти два стимула, произвести две оценки и одно сравнение между ними. Если же в одном пункте объединено четыре стимула, то количество оценок будет равно четырем (при этом испытуемому сложнее найти субъективную шкалу, адекватную всем четырем стимулам), а количество сравнений – шести. При этом всю информацию об оценках и сравнениях испытуемый вынужден держать в оперативной памяти. Следовательно, ему придется либо интенсивно использовать ее ресурсы оперативной памяти, либо принимать решения интуитивно. Объединение в один пункт более четырех стимулов в ипсативных методиках вынужденного выбора может привести к серьезным затруднениям для испытуемых.

На этапе компоновки теста необходимо следовать всем вышеприведенным правилам, применяемым к ипсативному пункту:

- каждый пункт должен объединять небольшое число шкал;
- все стимулы в рамках одного пункта должны обладать сходными психометрическими характеристиками и относиться к одной предметной области;
- каждая шкала должна объединяться в пунктах с каждой другой шкалой равное число раз.

При этом должны соблюдаться и требования к психометрическому качеству высказываний, равенства позитивных и негативных формулировок и т. д. (Closs, 1996).

**Исследовательский этап
разработки теста**

Исследовательский этап разработки ипсативной методики полностью включает действия, аналогичные тем, что производятся при разработке нормативной методики. Однако при этом существуют дополнительные факторы, которые следует контролировать. Например, в результатах ипсативной методики может проявиться систематическая ошибка, связанная с неоптимальной компоновкой стимулов или высказываний в пункты. Такая ошибка выявляется следующим образом: генерируется достаточно большой искусственный массив случайных ответов и проводится анализ интеркорреляционной матрицы результатов – все корреляции должны быть в среднем приближены к значению –

$$r_i = 1/(m - 1), \quad (4)$$

где m – число измеряемых шкал.

В том случае, если какая-то шкала коррелирует с другой сильнее, чем с остальными, то это, возможно, означает, что они чаще, чем следует, совмещаются в пунктах методики (Meade, 2004).

Также существует более проблемный тип систематических ошибок измерения, связанный с объединением шкал в пунктах. Даже если в одном пункте объединены стимулы, например а) с одинаковыми психометрическими характеристиками; б) не имеющие значительных факторных весов в одних и тех же шкалах; в) относящихся к одной предметной области и т. д., то характеристики этих двух стимулов в качестве целостного пункта все-таки остаются неизвестными (Stark, Chernyshenko, 2005). Как следствие, ипсативные инструменты при разработке требуют более глубоких и детальных пилотажных исследований уже скомпонованных версий методики, с необходимостью выявления в ходе эмпирических испытаний особенностей всех пунктов методики.

Особое внимание следует обратить на то, что концепция внутренней согласованности применима только к стимулам, составляющим шкалу, но не к пунктам (так как каждый пункт включает в себя несколько шкал). Концепция дискриминативности же применима и на уровне шкал, и на уровне пунктов – если большая часть испытуемых выбирает только одну из альтернатив пункта, то такой пункт следует пересмотреть или совсем заменить (Karpatschhof, Elkjaeer, 2000).

При оценке надежности и валидности ипсативных методик важно отметить, что недо-

пустима ссылка на относительно низкую надежность ипсативных методик как на неотъемлемую их особенность. Пытаясь вывести формулы для оценки надежности ипсативных методик, D. Bartram столкнулся со значительными проблемами. В частности, он выявил, что надежность ипсативной шкалы будет тем меньше, чем меньше шкал используется в ипсативной методике и чем выше положительные корреляции между этими шкалами в нормативном варианте (Bartram, 1996). Но в процессе разработки ипсативной методики должен быть обеспечен уровень надежности не ниже, чем у аналогичных нормативных инструментов. Для этого существует несколько методов – от простых (например, увеличение числа тестовых заданий) до более сложных, например, с применением психометрических подходов (Stark, Chernyshenko, 2005).

При проверке разных видов валидности, в частности, критериальной валидности, разработчик методики должен основываться на том, что описательные данные используемого критерия должны соответствовать ипсативным данным, то есть, так же отражать не абсолютные значения, а отношения. Например, если в качестве критерия выбрана аналогичная нормативная методика, то ее результаты должны быть прежде ипсатизированы (Berge, 1999).

Решение проблемы стандартизации ипсативной методики должно учитывать мнение, что разработка ипсативных методик менее затратна, чем разработка нормативных методик, так как ипсативным инструментам не требуется стандартизация в обычном смысле данного термина. Это положение вызывает возражение, поскольку стандартом и точкой отсчета в случае ипсативных методик является субъективность каждого конкретного испытуемого, что делает стандартизацию на уровне отдельных шкал невозможной. При этом должны быть ясно понимаемыми накладываемые этим обстоятельством ограничения.

Нормативные данные искусственно переводятся из исходной порядковой в равноинтервальную размерность шкалы посредством процедуры стандартизации. Ипсативные данные вследствие своих специфических свойств не могут быть подвергнуты такому преобразованию, что является еще одной из основных причин некорректности применения к ипсативным данным традиционных статистических процедур, применимых только к результатам равноинтервальных и более мощных шкал (Berge, 1999; Тепоруг, 1988).

Этот факт значительно ограничивает возможности применения ипсативных методик во многих сферах, в первую очередь, в исследовательской сфере и сфере психологической диагностики групп испытуемых. Однако малая «мощность» шкал ипсативных данных не ограничивает их применение для исследования единичных случаев, диагностики в рамках консультационной и клинической практик, самопознания (то есть в тех сферах, в которых ипсативные методы распространены исторически). Вопрос использования ипсативных методик в рамках профессионального отбора остается спорным, несмотря на распространенность такой практики (Vasilopoulos, Cucina, Dyomina, Morewitz, Reilly, 2006).

Выводы

В целом, разработка ипсативной методики является значительно более трудоемкой задачей, чем разработка аналогичной нормативной методики. Разработка и корректное практическое применение ипсативных методик, в частности, методик с ответами в формате вынужденного выбора, требует учета широкого спектра специфичных для ипсативных данных факторов. При этом выявляемые по сравнению с нормативными данными ограничения в статистическом анализе (являющиеся следствием уменьшения степеней свободы на одну) сочетаются с особыми требованиями к измеряемому конструкту, которые связаны с операциональной зависимостью ипсативных данных. Такие факторы значительно ограничивают возможности практического применения ипсативных методик, однако у них есть важные практические преимущества, например, сопротивляемость «обману» со стороны испытуемого, более детальное выявление структуры предмета диагностики и т. п. Указанные преимущества можно эффективно реализовать только при условии контроля в процессе разработки специфических для ипсативных методик факторов, описанных в настоящей статье.

Литература/References

1. Батури́н Н.А., Мельникова Н.Н. Технология разработки тестов: часть I. Вестник ЮУрГУ. Серия «Психология». 2009. Вып. 6. С. 4–14. [Baturin N.A., Melnikova N.N. Technology of tests development: Part I. *Bulletin of the South Ural State University. Ser. Psychology*, 2009, no. 6, pp. 4–14. (in Russ.)]
2. Батури́н Н.А., Мельникова Н.Н. Технология разработки тестов: часть II. Вестник ЮУрГУ. Серия «Психология». 2009. Вып. 7. С. 4–14. [Baturin N.A., Melnikova N.N. Technology of tests development: Part II]. *Bulletin of the South Ural State University. Ser. Psychology*, 2009, no. 7, pp. 4–14. (in Russ.)]
3. Батури́н Н.А., Мельникова Н.Н. Технология разработки тестов: часть III. Вестник ЮУрГУ. Серия «Психология». 2010. Вып. 8. С. 4–14. [Baturin N.A., Melnikova N.N. Technology of tests development: Part III]. *Bulletin of the South Ural State University. Ser. Psychology*, 2010, no. 8, pp. 4–14. (in Russ.)]
4. Батури́н Н.А., Мельникова Н.Н. Технология разработки тестов: часть IV. Вестник ЮУрГУ. Серия «Психология». 2010. Вып. 11. С. 13–28. [Baturin N.A., Melnikova N.N. Technology of tests development: Part IV]. *Bulletin of the South Ural State University. Ser. Psychology*, 2010, no. 11, pp. 13–28. (in Russ.)]
5. Батури́н Н.А., Мельникова Н.Н. Технология разработки тестов: часть V. Вестник ЮУрГУ. Серия «Психология». 2011. Вып. 12. С. 4–14. [Baturin N.A., Melnikova N.N. Technology of tests development: Part V]. *Bulletin of the South Ural State University. Ser. Psychology*, 2011, no. 12, pp. 4–14. (in Russ.)]
6. Ackerman T.A. Multidimensional Item Response Theory Modeling. In A. Maydeu-Olivares & J.J. McArdle. (Eds.), *Contemporary Psychometrics. Mahwah, NJ: Lawrence Erlbaum*, 2005, pp. 3–26.
7. Bartram D. The relationship between ipsatized and normative measures of personality. *Journal of Occupational and Organizational Psychology*, 1996, no. 69, pp. 25–39.
8. Bartram D. Increasing validity with forced-choice criterion measurement formats. *International Journal of Selection and Assessment*, 2007, no. 15, pp. 263–272.
9. Brown A. & Bartram, D. Doing less, but getting more: Improving Forced-Choice Measures with IRT. *Paper presented at the 24th Annual conference of the Society for Industrial and Organizational Psychology*, New Orleans, April 2–4, 2009.
10. Cattell R.B. Psychological measurement: Normative, ipsative, interactive. *Psychological Review*, 1944, no. 51, pp. 292–303.
11. Clemans W.V. An analytical and empirical examination of some properties of ipsative measures. *Psychometric Monographs*, 1966, pp. 14–28.
12. Closs S.J. On the factoring and interpretation of ipsative data. *Journal of Occupational and Organizational Psychology*, 1996, no. 69, pp. 41–47.

13. Cornwell J.M., & Dunlap W.P. On the questionable soundness of factoring ipsative data: A response to Saville & Willson (1991). *Journal of Occupational and Organizational Psychology*, 1994, no. 67, pp. 89–100.
14. Deaton W.L., Glasnapp D.R. & Poggio J.P. Effects of item characteristics on psychometric properties of forced-choice scales. *Educational and Psychological Measurement*, 1980, no. 40, pp. 599–610.
15. Hermelin E. and Robertson I.T. A Critique and Standardization of Meta-Analytic Validity Coefficients in Personnel Selection. *Journal of Occupational and Organizational Psychology*, 2001, no. 74, pp. 253–277.
16. Hicks L.E. Some properties of ipsative, normative, and forced-choice normative measures. *Psychological Bulletin*, 1970, no. 74, pp. 167–184.
17. Hough L.M. Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance*, 1998, no. 11, pp. 209–244.
18. Karpatschov B. and Elkjaer H.K. Yet the Bumblebee Flies: The reliability of ipsative scores examined by empirical data and a simulation study. *Research Report no 1. Department Psychology, University of Copenhagen*, 2000.
19. Martin B.A., Bowen C.C. & Hunt S.T. How effective are people at faking on personality questionnaires? *Personality and Individual Differences*, 2002, no. 32, pp. 247–256.
20. McCloy R., Heggestad E., Reeve C.A. Silk Purse From the Sow's Ear: Retrieving Normative Information From Multidimensional Forced-Choice Items. *Organizational Research Methods*, 2005, no. 8, pp. 222–248.
21. Meade, A. Psychometric Problems and Issues Involved with Creating and Using Ipsative Measures for Selection. *Journal of Occupational and Organizational Psychology*, 2004, no. 77, pp. 531–552.
22. Stark S., Chernyshenko O. & Drasgow F. An IRT approach to constructing and scoring pairwise preference items involving stimuli on different dimensions: The Multi-Unidimensional Pairwise-Preference Model. *Applied Psychological Measurement*, 2005, no. 29, pp. 184–203.
23. Berge T. A legitimate case of component analysis of ipsative measures, and partialling the mean as an alternative to ipsatization. *Multivariate Behavioral Research*, 1999, no. 34, pp. 89–102.
24. Tenopyr M.L. Artifactual reliability of forced-choice scales. *Journal of Applied Psychology*, 1988, no. 73, pp. 749–751.
25. Vasilopoulos N.L., Cucina J.M., Dyomina N.V., Morewitz C.L., & Reilly R.R. Forced-choice personality tests: A measure of personality and cognitive ability? *Human Performance*, 2006, no. 19, pp. 175–199.

Иванов Антон Евгеньевич, аспирант кафедры «Психологическая диагностика и консультирование», Южно-Уральский государственный университет (Челябинск), ivanovae36@bk.ru.

Поступила в редакцию 26 августа 2014 г.

**Bulletin of the South Ural State University
Series "Psychology"
2014, vol. 7, no. 3, pp. 14–21**

IPSATIVE PERSONALITY TESTS WITH ITEMS IN FORCED CHOICE FORMAT: FEATURES OF DEVELOPMENT

A.E. Ivanov, South Ural State University, Chelyabinsk, Russian Federation, ivanovae36@bk.ru.

Interest of study possible applications of ipsative data was revived in psychodiagnostics field in the past decade. In particular, the possibility of developing personality tests with items in forced choice format for professional selection. In process of creating this kind of tests it is necessary to consider some specific factors that do not affect the process of development of normative tests. This article contains results of the review of foreign literature on the subject of development and use of ipsative tests in comparison with normative, also description of specificity of both ipsative data and tests with items in forced choice format.

Keywords: ipsative test, ipsative data, ipsative scale, forced choice format, development of psychodiagnostic test.

Received 26 August 2014