

Психодиагностика

УДК 159.9.072

ИПСАТИВНЫЕ ЛИЧНОСТНЫЕ ТЕСТЫ ВЫНУЖДЕННОГО ВЫБОРА: АНАЛИЗ ЗАРУБЕЖНОГО ОПЫТА

Н.А. Батурин, А.Е. Иванов

В последнее десятилетие в зарубежной психодиагностике возрождается интерес к изучению ипсативного метода тестирования, в частности, к тестам формата вынужденного выбора. Приводятся результаты обзора иностранной литературы по данной тематике, рассмотрена история исследования ипсативного метода, описана специфика как ипсативных данных, так и ипсативных тестов вынужденного выбора. Проведен обзор преимуществ и ограничений ипсативных тестов вынужденного выбора, описываются последние исследования данной тематики.

Ключевые слова: ипсативный тест, ипсативные данные, ипсативные шкалы, формат вынужденного выбора, внутренняя зависимость, Multi-Unidimensional Pairwise Preference Model (MUPP).

Понятие ипсативного оценивания

Слово «ипсативный» имеет латинский корень «ipse», означающее «сам (лично)». В теории психологического измерения термин «ипсативный» используется для обозначения тех случаев, когда результаты диагностики каждого испытуемого служат измерительным эталоном для самих себя [2]. Иными словами, при анализе результатов тестирования применяются не внешние по отношению к испытуемому измерительные эталоны (нормы), а система субъективных сравнений и выборов конкретного испытуемого.

Автор термина R. Cattell внес в психологию понятие ипсативного шкалирования для разграничения субъективных и объективных эталонов (стандартов), относительно которых могут калиброваться психометрические инструменты [7]. Основным различием между ипсативным и нормативным шкалированием является выбор точки отсчета. От этого различия R. Cattell отталкивался при систематизации психодиагностических шкал. Вопрос, который он формулирует при этом, был следующим: «Какие эталоны можно использовать при выполнении психологического измерения?» [7, с. 292]. Ипсативное оценивание – это подход к оценке и интерпретации измеряемых показателей, основывающийся на установлении их отношения к выраженности других, с ними связанных. В таком случае каждый из диагностируемых показателей выра-

жается не в абсолютных значениях, а в форме доли относительно степени выраженности других показателей того же человека. Ипсативное оценивание является одной из альтернатив по отношению к нормативному оцениванию, при котором индивидуальный показатель сопоставляется с потенциальными статистическими характеристиками генеральной совокупности. Если в случае использования нормативной информации исследователь имеет дело с единицами измерения, отражающими выраженность некоего психологического феномена в определенной группе испытуемых, то, в случае ипсативной информации, исследователь имеет дело с единицами измерения, отражающими соотношение степени выраженности некоего психологического феномена по отношению к другим, измеренным данным инструментом, феноменам [7, 17]. В случае тестов диагностики личностных свойств ипсативные данные демонстрируют внутриличностное соотношение показателей и позволяют оценить индивидуально-типологическое своеобразие респондента [1].

Для ипсативного оценивания подходят только те области измерений, в которых респонденты способны осмысленно выбирать между ее альтернативами или ранжировать их в порядке предпочтения. Следовательно, для ипсативного оценивания не подходят области измерения, не связанные с системой предпочтений испытуемого. С другой стороны, ипса-

тивный формат подходит для оценки мотивационной направленности, личностных предпочтений, ценностей и других конструктов, в которых соотношение измеряемых характеристик имеет не меньшее значение, чем их абсолютная выраженность [10, 19, 20].

Первое оформление ипсативного подхода в психодиагностическом методе было сделано основателем Q-методологии В. Stephenson в 1935 г. В сороковые и пятидесятые годы двадцатого века в западной академической психологии проводились активные дебаты по поводу возможностей применения Q-методологии. Но, в конце концов эта методология получила более широкое применение в области социальных наук (политологии, социологии, маркетинге), а в психологии использовалась только Q-сортировка как отдельный метод, иногда применяемый и по сей день. Тем не менее, результаты Q-сортировки в психологии обычно подвергаются анализу в рамках традиционной для классической теории тестов R-методологии.

Интерес к ипсативному подходу в западной науке возродился в период 70-80 годов XX века, но этот интерес был связан скорее с преодолением психометрических ограничений ипсативных данных, без внимания к принципиальным различиям нормативной и ипсативной информации [12, 17]. В конце прошлого века было создано много ипсативных методик, большая часть которых основывалась на формате ответов в виде вынужденного выбора [4, 5, 13]. Эта тенденция чисто практического интереса к ипсативному подходу имела результатом то, что на сегодняшний день большинство исследователей определяют данные как «ипсативные», если «сумма сырых показателей, полученных по всем измеряемым шкалам, одинакова для всех испытуемых» [5].

Виды ипсативных данных

Разные авторы предлагают различающиеся классификации ипсативных данных, но наиболее распространено выделение трех видов:

- ипсатизированные данные AID (Additive Ipsative Data);
- композиционные ипсативные данные MID (Multiplicative Ipsative Data);
- ранговые ипсативные данные OID (Ordinal Ipsative Data) [12, 20].

Для удобства описания примем обозначения: $y = (y_1, \dots, y_k)'$ будет $k \times 1$ доипсативным

вектором нормативных измерений и $x = (x_1, \dots, x_k)'$ будет соответствующим $k \times 1$ преобразованным или ипсатизированным вектором от x .

Ипсатизированные данные (AID). Этот вид данных получается путем трансформации исходных нормативных данных (ипсатизации) следующим образом:

$$x_{AID} = (I - 1(1'1)^{-1}1')y = Ay, \quad (1)$$

где I – единичная матрица $k \times k$, 1 – это $k \times 1$ единичный вектор, $A = I - 1(1'1)^{-1}1'$ и $-$ это $k \times k$ аддитивная матрица преобразования.

Как следует из формулы (1), x_{AID} является вектором значений отклонения, полученных путем вычитания индивидуального среднего $U_{\text{среднее}} = (1/k)(y_1 + \dots + y_k)$ из сырых данных в y . Из-за этого x_{AID} так же называют *deviational ipsative data* [8]. Очевидно, что x_{AID} является ипсативным вследствие того, что сумма его элементов равна нулю. Так как в психологическом исследовании x_{AID} всегда получаются путем преобразования исходной матрицы y , то исследователь может анализировать оба набора данных. Иными словами, исследователь должен оценить потенциальные преимущества и ограничения, накладываемые использованием x_{AID} по отношению к использованию y . Например, если исходные данные подвержены сильному влиянию систематической ошибки измерения, которая завышает средние значения в матрице данных, то использование x_{AID} является оправданным. Так, в кросскультурных исследованиях группы испытуемых могут давать значительно различающиеся результаты, но это различие может быть следствием не истинных различий между людьми, а сторонних факторов, связанных, например с языковыми и культурными особенностями. Специальная процедура использования AID была разработана M. W.-L. Cheung и W. Chang для уменьшения влияния систематической ошибки измерения на результаты межгруппового (кросскультурного) структурного корреляционного анализа.

В кросскультурных исследованиях личности часто используется метод контроля систематической ошибки измерения, связанный со смещением средних значений в матрицах данных разных культурных групп, который называется «внутри субъектной стандартизацией» (*within-subject standartization*). При внутри субъектной стандартизации сырые данные сначала подвергаются процедуре ипсатизации, а затем находится частное от их

деления на значение индивидуального стандартного отклонения, в результате чего данные всех испытуемых уравниваются по среднему значению ($x_{\text{средн}} = 0$). Эта процедура математически может быть выражена следующим образом:

$$\begin{aligned} x_{\text{WS}} &= \sqrt{k-1} A y (y' A y)^{-\frac{1}{2}} = \\ &= \sqrt{k-1} (y' A y)^{-\frac{1}{2}} x_{\text{AID}}. \end{aligned} \quad (2)$$

Из приведенной формулы следует, что x_{WS} является взвешенным вариантом AID. Иначе говоря, если индивидуальные вариации между переменными схожи друг с другом, то x_{WS} и x_{AID} будут иметь схожие психометрические характеристики. Схожесть между x_{WS} и x_{AID} эмпирически зависит от специфики измерений и будет различаться в разных исследованиях [8].

Композиционные ипсативные данные (MID). Этот вид данных получается в том случае, если исходные данные представляют собой процентные доли от общего целого. Такого рода данные являются ипсативными, поскольку сумма всех переменных всегда равна 100 %. В связи с этим эти данные иногда обозначают как составные ипсативные данные (compositional ipsative data). Одним из преимуществ использования MID является то, что они позволяют сгладить влияние внутреннего напряжения испытуемого в случаях, когда затрагивается личная или конфиденциальная информация.

В отличие от AID, MID могут являться результатом использования определенного дизайна психодиагностического исследования. Тем не менее, если обозначить доипсативный вектор, из которого получены ипсативные данные, как y , то ипсативное преобразование будет иметь следующий вид:

$$x_{\text{MID}} = (1' y)^{-1} y \quad (3)$$

где x_{MID} это $k \times 1$ вектор MID.

Как следует из формулы (3), x_{MID} является мерой пропорции каждой переменной в векторе y по отношению к целому $1' y = \sum u_i$ для каждого испытуемого. Следует отметить, что x_{MID} останется неизменным в том случае, если y подвергнется влиянию масштабного коэффициента c (cy).

Ранговые ипсативные данные (OID) могут быть получены в том случае, если различные объекты ранжируются испытуемыми по определенной системе. Техника ранжиро-

вания применяется в тех случаях, когда оценка испытуемым альтернатив при помощи порядковых, интервальных и абсолютных шкал затруднена или невозможна. В случае, если число оцениваемых объектов невелико, то ранжирование является более простым и быстрым способом оценки, чем приписывание числовых значений.

Так же как и MID, OID могут являться результатом использования определенного дизайна психодиагностического исследования. Ипсативное преобразование, аналогичное приведенному в формуле (3), будет иметь следующий вид:

$$x_{\text{OID}} = F(y), \quad (4)$$

где $F(\cdot)$ – вектор-функция, по которой вычисляются ранги y .

При этом x_{OID} останется неизменным, даже если исходный вектор y будет подвержен воздействию масштабного и/или добавочного коэффициента.

Если исследователю доступны нормативные сырые данные, то они не должны быть ипсативированы без веских причин. Ипсатизация должна проводиться только в том случае, если существуют доказательства того, что исходные данные подвержены влиянию систематической ошибки измерения, равномерно искажающей эмпирические показатели относительно истинных. Там не менее, преобразуя нормативные данные в ипсативный формат, исследователь сталкивается с проблемой невозможности использования стандартных статистических методов. Согласно L.E. Hicks [12], использование ипсативных данных вместо их нормативных аналогов требует предварительного доказательства того, что:

- а) нормативные данные подвержены значительному влиянию систематической ошибки измерения;
- б) ошибка измерения снижает валидность в рамках данного исследования;
- в) ипсативный формат данных позволяет значительно уменьшить влияние систематической ошибки измерения и увеличить валидность диагностики эффективней, чем ипсативные методы контроля ошибки измерения.

На практике наиболее ярким проявлением ипсативного измерения является так называемое «самоипсативное шкалирование» (self-ipsative scaling), при котором формат ответов так или иначе требует от респондента распределить ресурсы, представленные определенным набором пунктов. Такого рода подсчет

показателей означает, что если некто получает высокий показатель по какой-то одной ипсативной переменной, он должен получить пропорционально низкие показатели по другим ипсативным переменным. В этом случае создается тенденция тяготения интеркорреляций между переменными к их отрицательным значениям.

В области диагностики свойств личности ипсативные данные чаще всего получают при помощи опросников с ответами в формате вынужденного выбора (FCID), в которых испытуемые должны выбрать из списка утверждения, которые «больше всего подходят» и «меньше всего подходят» им. В последние два десятилетия психологические исследования по вопросам, связанным с ипсативными данными, в основном касаются именно вопросов использования FCID, поскольку данный формат имеет потенциал к сокращению определенного рода систематических ошибок измерения, имеющих значение при личностной диагностике в рамках отбора персонала. В качестве примера популярной методики, основанной на FCID, можно назвать «Систему Томаса».

Специфика ипсативных тестов формата вынужденного выбора

Следует отметить, что сторонники использования ипсативных данных в психодиагностике сами не считают ипсативную информацию более удобной и адекватной изучаемым конструктам, чем нормативная информация, но уверены в существовании некоторых значительных преимуществ формата вынужденного выбора (FCID) над другими форматами тестовых пунктов [3, 11, 15]. Сторонники использования FCID приводят аргументы о лучшей сопротивляемости данного формата систематическим ошибкам, связанным со стилем ответа испытуемого, а так же большей стойкостью ипсативных методик к намеренной фальсификации результатов (faking). Хотя формат вынужденного выбора, несомненно, исключает наиболее простой способ фальсификации результатов (выставление высоких оценок по всем положительным характеристикам), он не запрещает испытуемому отвечать наиболее контекстуально приемлемым путем [12, 14].

Тем не менее, обман ипсативного теста требует от испытуемого больших затрат сил и высокого уровня когнитивных способностей, чтобы сформулировать корректные гипотезы

о том, какая из предложенных альтернатив характеризует его более положительно.

С другой стороны, получаемые при помощи FCID данные являются ранговыми ипсативными данными, что в сочетании с классической схемой разделения содержания опросника на пункты приводит к возникновению так называемой внутренней зависимости (internal interdependence). При использовании FCID получаемые данные подвержены влиянию двух типов внутренней зависимости.

Первый тип внутренней зависимости иногда обозначается как ковариационная зависимость, поскольку проявляется в ковариационной матрице ипсативных данных и, как следствие, в матрице интеркорреляций. Главной математической причиной специфической взаимосвязи ипсативных шкал является то, что они должны давать в сумме константный результат. В целом, увеличение количества конструктов уменьшает уровень ковариационной зависимости между ними. Точнее, с увеличением количества шкал будут уменьшаться чисто математические ограничения, накладываемые ипсативным форматом на ковариационную матрицу, и средняя ковариация между шкалами будет уменьшаться, стремясь от -1 к нулю. Тем не менее, корреляции между ипсативными факторами всегда отрицательны и средняя их величина равна $-1/(m-1)$, где m – это число шкал, даже если при использовании нормативных методов те же шкалы составляют ортогональные факторы.

Второй тип внутренней зависимости называется «внутренней или операциональной зависимостью на уровне пункта» (item-level interdependence). Обычно тест вынужденного выбора представляет собой набор пунктов, состоящих из сгруппированных утверждений. Испытуемый должен ранжировать утверждения в рамках каждого пункта или выбрать утверждение, которое «больше всего мне подходит» и «меньше всего мне подходит». При заполнении методики вынужденного выбора испытуемые вовлекаются в когнитивный процесс выбора [11]. Именно с этим процессом связан второй тип внутренней зависимости методик вынужденного выбора. В действительности, на выбор испытуемого влияет большее число факторов. Если представить данный выбор в виде формулы, аналогичной основной формуле классической теории тестов, то данная формула будет содержать не три переменные (как в случае нормативной

методики), а, как минимум, восемь переменных [17].

Выбор испытуемым одного из утверждений ипсативного пункта зависит не только от истинной выраженности латентной черты, измеряемой выбранным утверждением, но и от латентного уровня выраженности всех остальных черт, измеряемых данным пунктом, а также ошибок измерения, связанных со всеми включенными утверждениями, их «трудностью», уникальным стилем принятия решений, характерными для данного испытуемого. При этом следует отметить, что данный тип внутренней зависимости не присущ некоторым видам ипсативных данных, например, аддитивным ипсативным данным (AID), в которых исходная информация имеет нормативную природу.

Эти два типа внутренней зависимости напрямую определяют свойства ковариационной матрицы ипсативных данных, но при этом и дополнительные факторы также могут влиять на уровень внутренней зависимости [17]. Например, стратегия группировки утверждений в пункты опросника оказывает существенное влияние на свойства получаемых ипсативных данных. Оптимальным вариантом является схема, в которой каждая шкала сочетается во всех пунктах со всеми другими шкалами равное число раз. Если определенная шкала сочетается во всех пунктах чаще с какой-либо одной шкалой, чем с другими, то корреляции между этими шкалами будут тяготеть к большим отрицательным значениям, чем корреляции с другими шкалами. Более того, наблюдаемые значения шкал, группируемых чаще, чем другие, будут сильнее подвержены влиянию «истинных значений» и ошибок измерения друг друга.

Из-за сложных параметров пунктов шкальные оценки опросников вынужденного выбора представляют собой плохо контролируемое смешение различных факторов, как психометрических на уровне влияния пункта, так и скрытых когнитивных и личностных на уровне влияния испытуемого. В связи с этим ведется дискуссия относительно того, правомерно ли использовать ипсативные данные для пошкального сравнения результатов разных испытуемых [6].

Практическое применение тестов формата вынужденного выбора

С 1991 года среди специалистов англоязычных стран ведется активная дискуссия о

правомерности применения тестов формата вынужденного выбора. Основными пунктами этой дискуссии являются:

- соотношение психометрических характеристик ипсативных и нормативных данных;
- способность ипсативных методик нивелировать влияние сознательной фальсификации результатов диагностики;
- возможности получения при помощи FCID нормативных данных;
- и правомерность использования ипсативных данных для сравнения испытуемых друг с другом.

P. Saville и E. Willson показали, что искусственно сгенерированные данные OPQ32 по тесту CM7¹ дают корреляции между одними и теми же шкалами в разных форматах от 0,65 до 0,86 (в среднем 0,78). Авторы исследования полагают, что полученные корреляции достаточно высоки, чтобы признать результаты ипсативной и нормативной частей методики равнозначными. Они также утверждают, что причиной более низкой согласованности данных по шкалам вынужденного выбора являются ошибки измерения, присутствующие нормативным шкалам данного опросника, увеличивающие корреляцию между шкалами, а именно: тенденция к усреднению ответов, социальная желательность и тенденция выбора положительных ответов [17]. Против такой интерпретации возражает A. Meade, так как значение $r = 0,8$ объясняет 64 % дисперсии (36 % наблюдений все еще ничем не объясняются). S.J. Closs разделяет точку зрения A. Meade и возражает против использования ипсативных методик для межличностного сравнения, но не возражает против использования такого рода методик в сфере отбора персонала. Он также критикует P. Saville и E. Willson, поскольку они полагаются на теоретические и искусственные процедуры генерации данных [9]. В противовес этой точке зрения В. Karpatschhof & Н.К. Elkjaer утверждают, что ипсатизированные и ипсативные измерения могут быть использованы для межличностного сравнения [13]. Они указывают на то, что оба формата измерения направлены на оценку одних и тех же поведен-

¹ CM-7 представляет собой тест из двух измерений, различающихся только форматами пунктов: одна его часть выполнена в нормативном формате шкал Лайкерта, а другая в формате вынужденного выбора.

ческих феноменов. Такой вывод был сделан и в ранних исследованиях P. Saville и E. Willson, G. Matthews и K. Oddy [17], а также высказывался D. Bartram [4].

Так как ипсативные и нормативные измерения теоретически являются равноценными мерами соответствующих «истинных баллов», B. Karpatschof & H.K. Elkjaer, P. Saville и E. Willson, G. Mattie и K. Oddy, D. Bartram полагают, что ипсативные измерения могут на уровне популяции выявлять искаженные, но все же достаточно точные меры индивидуальных различий. Их критики приводят контраргументы, указывающие на то, что ипсативные тесты имеют слишком сложную для анализа ошибку измерения и значительные статистические несовершенства, которые не присущи нормативным тестам, что делает ипсативные тесты непригодными для межиндивидуального сравнения [9, 12, 19].

В свою очередь, R. McCloy доказывает, что ипсативные данные могут быть преобразованы в нормативные с использованием методологии IRT и модели развертывания Кумбса («Coombs unfolding models»), а затем использованы для сравнения результатов испытуемых между собой [16]. R. McCloy возражает против использования ипсативных методик в контексте задач профотбора. По его мнению, предложенная им модель извлечения нормативной информации «является только первой демонстрацией того, что ипсативные данные могут предоставлять информацию об абсолютной (в противовес относительной) выраженности изучаемых феноменов» [16, с. 228]. Ранее J.M. Cornwell и W.P. Dunlap обосновали противоположные этому выводы, согласно которым трансформация может производиться только от нормативных данных к ипсативным, так как это путь потери информации. Однако в настоящее время стало очевидно, что их аргументация имеет силу по отношению не ко всем способам получения ипсативных данных [11]. Теория R. McCloy послужила одной из основ модели MUPP, позволяющей на практике разрабатывать методики вынужденного выбора, из которых может быть извлечена информация нормативного характера [18].

Активно обсуждается надежность ипсативных методик. M.L. Tenopir & C.E. Johnson et al. [20] полагают, что надежность ипсативных методик искусственно завышается, хотя P. Saville & E. Willson [18], основываясь на уже упомянутом аргументе о большем коли-

честве источников систематических ошибок, показали, что внутренняя согласованность ипсативных инструментов ниже, чем у аналогичных по формальным характеристикам нормативных. Это объясняется большим количеством дополнительных факторов ошибки измерения, вызванных внутренней зависимостью. D. Bartram [4] и B. Karpatschof & H.K. Elkjaer [14] возражают, что данные систематические ошибки имеют место, но не оказывают существенного влияния. Если ипсативные инструменты ненадежны, то и оценка валидности не имеет значения. Существуют эмпирические исследования, указывающие на то, что частично ипсативные шкалы обладают сходным с нормативными шкалами уровнем валидности [5, 6, 18]. Но при этом, как показала практика, возможен контроль вышеупомянутых источников, позволяющий значительно улучшить показатель надежности опросников вынужденного выбора [6].

Уровень валидности ипсативного инструментария, так же как и надежность, по мнению N. Christensen et al. и A. Meade, находится под вопросом из-за спорности соответствия изучаемых конструктов ипсативному формату, а также более сложных процессов сравнения и принятия решения при выполнении ипсативных методик [17]. Существенное влияние на валидность так же оказывает связь между шкалами ипсативной методики, возникающая из-за включения стимулов из разных шкал в один пункт.

Одним из последних крупных исследований валидности применения ипсативных методик в ситуации психодиагностики личности в рамках профотбора является метаанализ, выполненный D. Bartram [4]. В вышеуказанном метаанализе были собраны данные 11 исследований, 9 из которых были проведены в англоязычных странах (Великобритании, США, ЮАР), а остальные два – в Турции и Южной Корее. Общий объем выборки англоязычных стран составил 1578 испытуемых, в турецкой и корейской выборках – 364 и 503 соответственно. Все данные выборки включают испытуемых, тестиовавшихся в реальной ситуации отбора на работу, связанную с управленческой деятельностью. Эти исследования объединены использовавшимся инструментарием – опросниками OPQ32n и OPQ32i (SHL). Эти опросники представляют собой две версии одной методики, построенной на модели «большой восьмерки», используемой для оценки личности человека в рабо-

чем контексте. В заключение своей работы D. Bartram делает вывод, что ипсативные методики имеют ряд психометрических особенностей, их использование оправдано практикой, так как они более устойчивы со стороны попыток искажения результатов испытуемым и обладают большей валидностью [5].

Разработчики пришли к выводу, что в классическую теорию тестов не вписывается процесс принятия решения или выбора, свойственный ответу на ипсативный пункт. Им нужна была модель, позволяющая адекватно оценивать процесс выбора, в который включается испытуемый, способная улучшить качество результирующей информации.

В 2007 году S. Stark & O.S. Chernyshenko предложили модель, названную «Multi-Unidimensional Pairwise Preference Model» (MUPP). Суть данной модели заключается в адаптивном объединении в пары утверждений с известными характеристиками в рамках теории IRT [18]. В результате проведения всех необходимых сравнений (дальнейшие исследования выявили, что не обязательно производить все возможные сравнения, пригодным является более экономичный адаптивный циркулярный метод), у исследователя появляется возможность ранжировать стимулы методики по степени предпочтения испытуемым. Этот ранговый ряд в дальнейшем может быть использован для извлечения данных, пригодных для стандартизации, то есть по сути нормативных. Авторы метода в настоящее время пытаются применить MUPP к разработке многомерных методик вынужденного выбора, но пока им удалось достичь его адекватного применения только в случае небольшого числа измеряемых шкал (не более 10), так как количество сравнений, необходимых для оценки латентной черты таким способом, слишком велико. Дополнительно, для извлечения данных абсолютной природы, даже в многомерных методиках, сконструированных с использованием MUPP, необходим некоторый минимальный процент адаптивно создаваемых одномерных пунктов вынужденного выбора. Например, при создании десятишкальной методики вынужденного выбора необходимо минимум 10 % одношкальных пунктов. Существуют адаптивный и фиксированный варианты методик на основе MUPP, при этом адаптивный вариант позволяет получить несколько более точные данные, например, корреляции между известным порядковым рядом выраженности латентных черт и

измеренным составляют 0,82 (фиксированный вариант) и 0,94 (адаптивный вариант) для пятимерного теста из 100 пунктов [18].

A. Brown & D. Bartram использовали разработку S. Stark & O.S. Chernyshenko по применению IRT-моделей при разработке методик вынужденного выбора, но использовали более простую неадаптивную модель The Two-dimensional IRT Preference Model for Paired Comparisons, а в саму методику не включили одношкальные пункты. Разработчики новой версии OPQ32i поставили три основные задачи по ее модификации:

1. Уменьшить количество утверждений в каждом пункте с 4 до 3.

2. Сохранить имеющиеся психометрические характеристики методики.

3. Обосновать возможность получения абсолютных, а не относительных данных, при использовании методики.

С результатами данной модификации можно ознакомиться в соответствующей статье [6]. Здесь отметим только, что обоснование нормативности получаемых данных, приведенное A. Brown & D. Bartram представляется менее доказательным, чем аналогичные разработки R. McCloy, так как не переводит процедуру тестирования в адаптивную форму.

Выводы

Ипсативные психодиагностические данные, изначально анализируемые в рамках Q-методологии, в настоящий момент рассматриваются в рамках R-методологии. В связи с этим выходят на первый план некоторые особенности ипсативных данных, плохо вписывающиеся в рамки R-методологии, в частности, внутренняя зависимость.

Первоначально считалось, что ипсативные данные могут быть использованы только для внутрииндивидуального сравнения, но в последнее десятилетие многие исследователи ищут способ доказать пригодность ипсативных данных в таких сферах, как оценка и отбор персонала. Данные тенденции вызваны, в первую очередь, тем, что тесты формата вынужденного выбора, которые позволяют получить только ипсативные психодиагностические данные, согласно некоторым исследованиям, более устойчивы к сознательным и, в меньшей степени, бессознательным, искажениям результатов со стороны испытуемого. Таким образом, существует конфликт: с одной стороны тесты формата вынужденного выбора лучше подходят для сфер профотбора

благодаря сопротивляемости «обману» со стороны кандидатов, а с другой стороны тесты формата вынужденного выбора ипсативны, потому их результаты не могут служить надежной основой для сравнения и ранжирования испытуемых.

В настоящее время вышеуказанный конфликт решается путем модификации тестов формата вынужденного выбора таким образом, чтобы получаемые с их помощью результаты могли быть приравнены к нормативным данным. Существует два основных подхода к решению этой задачи: психометрический и с использованием адаптивного тестирования.

Суть психометрического подхода заключается в применении максимально возможного контроля над факторами внутренней зависимости и приближении психометрических характеристик тестов вынужденного выбора к показателям аналогичных нормативных тестов. Основным сторонником данного подхода является D. Bartram и компания SHL, которая разработала ипсативный вариант своей методики OPQ32. Ипсативный вариант OPQ32 примерно в два раза длиннее, чем нормативный, и более трудозатратен при заполнении для испытуемого, но имеет психометрические характеристики, сравнимые с аналогичными характеристиками нормативного варианта OPQ32.

Применение IRT и адаптивного тестирования с использованием MUPP (использование в пунктах формата вынужденного выбора утверждений, относящихся в одной шкале) также позволяет извлечь нормативные данные.

Литература

1. Аксенова, Е.П. Ипсативный подход к оценке личности: возможности и ограничения / Е.П. Аксенова // Психологические исследования: электрон. науч. журн. – 2012. – Т. 5, № 24. – <http://psystudy.ru>
2. Бурлачук, Л.Ф. Словарь-справочник по психодиагностике. / Л.Ф. Бурлачук, С.М. Морозов – Киев: Наукова Думка, 1989. – 200 с.
3. Baron, H. Strengths and limitations of ipsative measurement / H. Baron // *Journal of Occupational and Organizational Psychology*. – 1996. – № 69. – P. 49–56.
4. Bartram, D. Increasing validity with forced-choice criterion measurement formats / D. Bartram // *International Journal of Selection and Assessment*. – 2007. – № 15. – P. 263–272.
5. Bartram, D. The relationship between ipsatized and normative measures of personality / D. Bartram // *Journal of Occupational and Organizational Psychology*. – 1996. – № 69. – P. 25–39.
6. Brown, A. Doing less, but getting more: Improving Forced-Choice Measures with IRT / A. Brown, D. Bartram // Paper presented at the 24th Annual conference of the Society for Industrial and Organizational Psychology. New Orleans, April 2–4, 2009.
7. Cattell, R.B. Psychological measurement: Normative, ipsative, interactive / R.B. Cattell // *Psychological Review*, 1944. – № 51. – P. 292–303.
8. Chan, W. The covariance structure analysis of ipsative data / W. Chan, P.M. Bentler // *Sociological Methods and Research*. – 1993. – № 22. – P. 214–247.
9. Closs, S.J. On the factoring and interpretation of ipsative data / S.J. Closs // *Journal of Occupational and Organizational Psychology*. – 1996. – № 69. – P. 41–47.
10. Dunlap, W.P. Factor analysis of ipsative measures / W.P. Dunlap, J.M. Cornwell // *Multivariate Behavioral Research*. – 1994. – № 29. – P. 115–126.
11. Forced-choice personality tests: A measure of personality and cognitive ability? / N.L. Vasilopoulos, J.M. Cucina, N.V. Dyomina et al. // *Human Performance*. – 2006. – № 19. – P. 175–199.
12. Hicks, L.E. Some properties of ipsative, normative and forced-choice normative measures / L.E. Hicks // *Psychological Bulletin*. – 1970. – № 74. – P. 167–184.
13. Jackson, D.N. The Impact of Faking on Employment Tests: Does forced-choice offer a solution? / D.N. Jackson, V.R. Wroblewski, M.C. Ashton // *Human Performance*. – 2000. – № 13. – P. 371–388.
14. Karpatschhof, B. Yet the Bumblebee Flies: The reliability of ipsative scores examined by empirical data and a simulation study / B. Karpatschhof, H.K. Eljaejer // Research Report no 1. Department Psychology, University of Copenhagen, 2000.
15. Martin, B.A. How effective are people at faking on personality questionnaires? / B.A. Martin, C.C. Bowen, S.T. Hunt // *Personality and Individual Differences*. – 2002. – № 32. – P. 247–256.
16. McCloy, R. A Silk Purse From the Sow's Ear: Retrieving Normative Information From Multidimensional Forced-Choice Items / R. McCloy, E. Heggstad, C. Reeve // *Organizational Research Methods*, 2005. – № 8. – P. 222–248.

17. Meade, A. *Psychometric Problems and Issues Involved with Creating and Using Ipsative Measures for Selection* / A. Meade // *Journal of Occupational and Organizational Psychology*. – 2004. – № 77. – P. 531–552.

18. Stark, S. *An IRT approach to constructing and scoring pairwise preference items involving stimuli on different dimensions: The Multi-Unidimensional Pairwise-Preference Model* / S. Stark, O. Chernyshenko, F. Drasgow // *Applied*

Psychological Measurement. – 2005. – № 29. – P. 184–203.

19. Ten Berge, J.M.F. *A legitimate case of component analysis of ipsative measures, and partialling the mean as an alternative to ipsatization* / J.M.F. Ten Berge // *Multivariate Behavioral Research*. – 1999. – № 34. – P. 89–102.

20. Tenopyr, M.L. *Artifactual reliability of forced-choice scales* / M.L. Tenopyr // *Journal of Applied Psychology*. – 1988. – № 73. – P. 749–751.

Батурин Николай Алексеевич, доктор психологических наук, профессор, заведующий кафедрой психодиагностики и консультирования, Южно-Уральский государственный университет, г. Челябинск, nikbat@list.ru

Иванов Антон Евгеньевич, аспирант кафедры психодиагностики и консультирования, Южно-Уральский государственный университет, г. Челябинск, ivanovae36@bk.ru

IPSATIVE PERSONALITY TESTS WITH ITEMS IN FORCED CHOICE FORMAT: ANALYSIS OF INTERNATIONAL EXPERIENCE

N.A. Baturin, South Ural State University, Chelyabinsk, Russian Federation, nikbat@list.ru

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

In the past decade, foreign psychodiagnostics revived interest in the study of ipsative testing method, in particular, of forced choice format tests. This article contains results of the review of foreign literature on the subject, reviews of the history of research on ipsative method, description of the specific of ipsative data and ipsative forced choice test. This is a review of the advantages and limitations of ipsative forced choice tests and description recent research on this topic.

Keywords: ipsative test, ipsative data, ipsative scale, forced choice format, internal interdependency, Multi-Unidimensional Pairwise Preference Model (MUPP).

Поступила в редакцию 19 августа 2013 г.