

ТЕХНОЛОГИЯ РАЗРАБОТКИ ТЕСТОВ: ЧАСТЬ IV

Н.А. Батурич, Н.Н. Мельникова

Статья является продолжением описания пошаговой технологии разработки тестов, изложение которой начато в предыдущих номерах журнала. В четвёртой части цикла статей описывается VI «интерпретационный» этап разработки теста и рассматриваются ключевые задачи этого этапа, связанные с разработкой схем интерпретации, подготовкой текстов интерпретаций и описанием различных моделей составления отчётов. Также в статье подробно обсуждаются три принципа, обеспечивающие качество интерпретаций и валидность диагностических отчётов: (1) ориентация на конкретные практические цели, (2) соблюдение границ содержания, (3) опора на эмпирические данные, полученные в ходе психометрической проверки методики.

Ключевые слова: разработка теста, интерпретация теста, диагностические отчёты.

Введение

В первой части цикла статей была представлена общая схема процесса разработки тестов, приведена последовательность этапов и шагов разработки и указаны основные задачи, решаемые на каждом из них [2]. Там же были описаны два начальных этапа: этап I – «организационный» и этап II – «содержательный», которые являются фундаментом создания любой методики, задавая стратегическую линию разработки и обеспечивая валидность методики на содержательном уровне. Во второй части цикла статей обсуждались следующие этапы: III – «подготовительный» и IV – «исследовательский», посвященные вопросам формулирования и отбора эффективных тестовых пунктов и проблемам проверки основных психометрических показателей тестов: надёжности и валидности [3]. В третьей части рассматривался V – «стандартизационный» этап [4]. Данная статья продолжает поэтапное описание процесса разработки тестов и посвящена вопросам интерпретации теста.

Этап VI. ИНТЕРПРЕТАЦИОННЫЙ Шаг 10. Разработка системы интерпретации результатов

Шестой «интерпретационный» этап разработки чрезвычайно важен для подготовки методики к её корректному применению. Неграмотная интерпретация полученных результатов может свести «на нет» все усилия разработчиков по достижению валидности теста, предпринятые на предыдущих этапах процесса. Проблема заключается в том, что в совре-

менной отечественной практике использования тестов ответственность за интерпретацию результатов во многом возложена на пользователя. Разработчики обычно не уделяют должного внимания проработке этапа интерпретации и ограничиваются кратким определением изучаемых конструктов и абстрактными описаниями «высоких» и «низких» баллов по отдельным шкалам. Однако для создания диагностического отчёта, который мог бы быть полезен для практики, такой информации недостаточно. Это особенно актуально для сложных многошкальных методик, где наиболее информативными являются комбинации показателей [6]. В отечественной практике мы привыкли к тому, что грамотность применения той или иной методики достигается преимущественно опытом её использования. В результате многократных (и многолетних) применений у пользователя формируется качество «экспертности» – во многом интуитивное «видение» того, как интерпретировать конкретную конфигурацию данных, какую стратегию анализа лучше выбрать, на что обратить внимание в первую очередь и т. д. [12]. Путём многократных проб и ошибок пользователь приходит к умению создавать валидные интерпретации и заключения, поскольку вынужден соотноситься с практикой и отвечает за результат.

Однако для пользователя длительность этого пути и сложность действий, связанных с интерпретацией и составлением диагностических отчётов, может быть значительно снижена усилиями разработчиков методики. Если

удаётся формализовать необходимые процедуры: прописать схемы анализа, составить удобные алгоритмы, а ещё лучше – автоматизировать ряд задач, передав их компьютеру, пользователь оказывается в более благоприятной ситуации, а методика становится лучше застрахованной от некорректного применения [6]. На VI этапе должна быть разработана система интерпретации, сопровождающаяся подробными указаниями для пользователя, регламентирующими и направляющими его действия при интерпретации результатов и составлении диагностических отчётов. В целом, на шестом «интерпретационном» этапе процесса создания методики перед разработчиками стоят 3 задачи: (1) разработка схем интерпретации, (2) подготовка текстов интерпретаций, (3) описание моделей составления отчётов.

Перед тем как рассмотреть особенности выполнения названных задач (в статье это будет сделано позже), важно выделить базовые принципы, на которые опирается такая работа.

Три принципа, обеспечивающие качество интерпретаций и валидность диагностических отчётов

Критерием оценки качества интерпретаций и построенных на их основе диагностических отчётов может служить соответствие следующим принципам:

- 1) *ориентация на конкретные практические цели;*
- 2) *соблюдение границ содержания;*
- 3) *опора на эмпирические данные, полученные в ходе психометрической проверки методики.*

Указанные принципы – это «три кита» хорошей интерпретации и условия, обеспечивающие валидность и практическую применимость диагностических отчётов. Поэтому они должны служить руководящими основами, направляющими деятельность разработчика на интерпретационном этапе. Рассмотрим эти принципы подробнее.

1. Ориентация на конкретные практические цели

Ни у кого не вызывает сомнения утверждение, что диагностика не является самоцелью. Полученная с помощью тестов информация нужна для решения каких-то практических задач. Например, если на некотором предприятии используется тест трудовой мо-

тивации, то, как правило, конечная цель диагностики не заключается в том, чтобы просто «посмотреть, какая у сотрудников мотивация». Реальные цели более прагматичны и более конкретны. Так, результаты исследования трудовой мотивации могут быть использованы для: а) разработки программы нематериального мотивирования сотрудников б) отбора кандидатов на определённые вакантные должности (например, на управленческие), в) прогноза успешности адаптации вновь принятых сотрудников и т.д.

Разные практические цели предполагают разные стратегии анализа результатов, разные ключевые единицы интерпретации и разные формы отчётов. Например, при использовании теста трудовой мотивации для целей отбора необходимо соотнесение полученных результатов с критериальными нормами, а при планировании программы мотивации более продуктивной будет интерпретация ведущих мотивов в свете репрезентативных групповых норм. При отборе кандидатов может быть полезным анализ типовых профилей, а для разработки программы мотивации – анализ тенденций по структурным единицам предприятия. Отчёт, полезный для принятия решения о приёме или отклонении кандидата, должен содержать оценку степени риска позитивного и негативного решения и описание переменных-модераторов, способных повлиять на прогнозируемый исход. А в отчёт, который будет использоваться для планирования мероприятий по мотивированию сотрудников, желательно включить тексты, поясняющие организационные факторы, отвечающие за те или иные мотивационные тенденции.

Пригодный для практики диагностический отчёт должен содержать ключевую информацию, важную для принятия тех или иных решений. А задача разработчика – предоставить пользователю *технологии извлечения из полученных результатов сведений, полезных для конкретной практической цели* [6]. Если методика многоцелевая (т.е. планируется её использование для нескольких практических целей), то для каждой цели предлагается своя схема анализа, свои тексты интерпретаций и своя форма отчётов.

Такой подход может показаться нестандартным (а точнее, непривычным для отечественной психодиагностики), но он вполне оправдан. Напомним, что с точки зрения практики (и современной психометрики) такое понятие, как некая «абстрактная» валид-

ность теста считается бессмысленным [1, 7, 10, 15]. Общая концептуальная валидность и интерпретация, делающая акцент на сущности изучаемого конструкта и построенная на детальном описании его особенностей, несомненно, представляет ценность для любого инструмента. Однако это прерогатива исследовательской методике. Для того чтобы инструмент мог быть эффективно использован на практике, таких сведений недостаточно. Практико-ориентированные интерпретации – ещё одно существенное отличие диагностической методике от исследовательской.

Ещё на I «организационном» этапе разработки (см. шаг 1 «Планирование проекта») ставится вопрос: для поддержки каких практических задач требуется разработать инструмент? Именно цели практики определяют «область валидности» теста, в соответствии с чем уже и планируются все процедуры её обеспечения и проверки. Напомним, что разные практические цели предполагают разные программы проверки валидности и разные модели стандартизации [2]. Если такие вопросы продуманы изначально, то на этапе интерпретации не возникает сложностей с тем, чтобы приспособить методику к конкретным практическим задачам.

Чёткая направленность интерпретации на конкретные практические цели застрахует тест ещё от одной серьёзной проблемы: тенденции к излишнему расширению области его применения. Простая истина заключается в том, что даже самый многоплановый психодиагностический инструмент (например, такой как 16-PF или Vig-5) не может быть универсальным. Поэтому с некоторым скепсисом приходится относиться к популярным попыткам использовать тот же 16-PF для всевозможных целей (профотбора, личностного консультирования, подбора команд и т. д.), если эти возможности инструмента специально не валидизированы.

При этом не исключается появление новых областей практического использования теста в процессе его эксплуатации. Однако повторимся ещё раз, эти новые возможности должны быть специально валидизированы и представлены в виде отдельной пользовательской стратегии. Например, существует специальная версия CPI, приспособленная для целей консультирования по вопросам развития лидерских качеств [18]. Для этих целей авторы предлагают отдельное руководство к CPI, где обозначены определённые комбинации

шкал, имеющие отношение к тем или иным характеристикам лидерства, и для них проведена специальная стандартизация.

Таким образом, у завершённого и готового к применению психодиагностического инструмента схемы интерпретаций и модели отчётов должны быть привязаны к конкретным практическим целям. В руководстве также полезно обозначить ограничения методике: указать, для каких практических целей методика не может применяться.

2. Соблюдение границ содержания

Смысл этого принципа заключается в том, что интерпретации не должны выходить за границы того содержания, которое изначально заложено в тест. На первый взгляд, этот принцип представляется очень естественным и простым. Однако на практике мы достаточно часто сталкиваемся с неадекватно расширенными толкованиями результатов, а иногда – и с явными несоответствиями текстов интерпретаций тому диагностическому материалу, на основе которого получены данные.

Это особенно актуально для методик, в основе которых лежат сложные, неоднозначные концепты, например такие, как жизнестойкость, эмоциональный интеллект, честность, толерантность и др. Часто тексты интерпретаций составляются только с опорой на название и выглядят, как абстрактная словарная статья, допускающая очень широкую содержательную трактовку. Например, жизнестойкость может рассматриваться и как аналог стрессоустойчивости, и как умение преодолевать трудности на пути достижения жизненных целей. А межличностную толерантность можно трактовать и как нечувствительность к определённым проявлениям другого человека, и как принятие, основанное на признании ценности другого. Хорошая интерпретация должна прояснять, какое именно концептуальное понимание заложено в тест автором. Иначе для пользователя создаётся искушение вольной трактовки результатов, разворота контекста в сторону, субъективно интересующую пользователя или удобную для заказчика. Из-за этого в процессе эксплуатации нередко происходит «смещение» содержания теста, что негативно влияет на его валидность.

Напомним, что ещё на II «содержательном» этапе процесса разработки теста перед авторами стояла задача определения концепта [2]. Явлению, для измерения которого предназначен тест, было дано чёткое определение,

отражающее концептуальную позицию автора, его понимание и специфическую трактовку. Именно эта трактовка и должна быть подчеркнута в текстах интерпретаций, чтобы избежать искажения содержания, которое может быть привнесено пользователем.

В целом, при грамотной организации процесса разработки теста ещё на его начальных стадиях были подробно описаны особенности содержательного поля и ключевой феноменологии, впоследствии заложенные в тест. Эта информация может быть весьма полезна при составлении текстов интерпретаций, поскольку позволяет чётче обозначить границы содержания, отвечая на вопросы: в каких областях реальности и посредством каких эмпирических феноменов свойство, измеряемое тестом, даёт о себе знать [2].

Если в качестве интерпретации автор предлагает лишь абстрактное описание концепта, то область валидности теста становится неопределённой, и соответственно снижается его практическая ценность. Некоторые искушённые в этом вопросе пользователи для более ясного вычленения и ограничения содержательного поля интерпретаций прибегают к «кустарному» приёму: они обращаются к конкретным пунктам теста, и это конкретное содержание закладывают в основу интерпретаций. Таким способом пользователи соблюдают требования содержательной валидности, стараясь «не попасть мимо» в своих заключениях. Однако задача чёткой прорисовки и ограничения содержательной области интерпретаций всё же должна быть задачей разработчика, а не вынужденной мерой пользователя.

В этом смысле в наиболее выгодной позиции находятся разработчики тестов достижений. Многие широко востребованные образцы таких тестов сопровождаются предметно-ориентированными нормами и предлагают интерпретацию, которая строится на весьма детальном описании уровней выполнения через изложение содержания тестовых единиц, соответствующих каждому уровню [14]. Иногда в отчётах в качестве иллюстраций даже приводятся таблицы с описанием содержания отдельных, наиболее репрезентативных для каждого субтеста пунктов. Проведённые фокус-групповые исследования, где участниками были лица, использующие результаты тестов достижений в своей работе, показали, что наиболее полезными и информативными для принятия практических решений считаются отчёты, предлагающие подробные описания и

объяснения содержательной стороны теста. По мнению участников фокус-групп, такие отчёты позволяют сориентироваться в том, что конкретно измерялось и направленно выделить сильные и слабые стороны каждого тестируемого [19].

Естественно, не всегда возможно достичь такой степени детализации содержания при интерпретации личностных методик. Однако при составлении текстов интерпретаций необходимо учитывать особенности того ситуативного поля, на материале которого конструировался тест. Так, например, при интерпретации теста, диагностирующего ответственность, важно указать, были ли использованы ситуации, связанные с профессиональной деятельностью, или же ответственность изучалась на материале ситуаций межличностных отношений, а может, – на материале стрессовых ситуаций. Заложенное в тест содержательное поле задаёт контекст интерпретаций и во многом определяет границы экстраполяции результатов. Поэтому такие сведения обязательно должны быть донесены до пользователя.

Особая ситуация складывается с интерпретацией шкал, полученных в результате факторизации. Такие шкалы редко имеют хорошо проработанные содержательные основы, поскольку многие из них выявляются только в ходе создания методики. Разработчики, имеющие опыт использования факторной модели, знают, как сложно дать корректное название найденному фактору, которое бы точно отражало сущность латентной переменной. Тем не менее, когда название всё же выбрано, его почему-то воспринимают как окончательную истину, и интерпретацию составляют уже исходя из названия, часто серьёзно отклоняясь от реального эмпирического наполнения фактора. Но для интерпретатора сложность состоит ещё и в том, что даже опора на содержание пунктов, получивших наибольший вес в факторной структуре, не всегда позволяет уловить суть латентной переменной. Поэтому любой вновь полученный фактор должен быть специально валидизирован (см. далее принцип эмпирического подтверждения интерпретаций), после чего имеет смысл вновь задуматься о его наименовании.

В целом, содержание интерпретаций должно быть привязано к конкретному инструменту, а не просто к соответствующему абстрактному концепту. Поэтому важно, чтобы при подготовке текстов интерпретаций разработчик учитывал два момента: а) особенности

авторской трактовки явления, б) содержательное поле и ключевую феноменологию, на которых построен тест.

3. Опора на эмпирические данные, полученные в ходе психометрической проверки методики

В стандартах AERA, APA & NCME зафиксировано, что любая интерпретация должна иметь как теоретические основы, так и эмпирические подтверждения [13]. Мы не можем вносить в интерпретацию того содержания, которое не подтверждено эмпирически. Например, как уже было указано выше, чтобы корректно интерпретировать некий фактор, полученный в ходе разработки теста, он должен быть специально валидизирован.

Основным источником эмпирической информации при подготовке содержания интерпретаций выступают данные, полученные в ходе психометрической проверки и стандартизации методики. При добросовестно организованной разработке таких данных может быть достаточно много. Однако для разработчика существуют 4 «скользких» момента при подготовке интерпретаций. Они связаны

- с интерпретацией *паттернов, конфигураций показателей* (например, профилей),
- интерпретацией *производных показателей* (например таких, как индексы, коэффициенты),
- интерпретацией *шкал, полученных в результате факторизации*,
- включением в тексты интерпретаций *ссылок на другие переменные, связанные с диагностируемой*.

Кратко остановимся на каждом из этих пунктов.

Выше уже говорилось, что наиболее полезная для практики информация по большей части извлекается из определённых *комбинаций показателей*. Многие разработчики сопровождают интерпретации тестовых батарей объяснениями того, что значат те или иные сочетания показателей, а в руководствах к многошкальным методикам предлагают интерпретацию типовых профилей. Отдельные профили могут рекомендоваться как критерии отбора (например, на определённую должность) или служить показателями некоторого диагноза.

Подобные стратегии интерпретации могут быть продуктивными и весьма удобными для пользователя. Основная проблема здесь

заключается в том, что при разработке методик традиционно *валидизируются отдельные шкалы*, в то время как *интерпретируются в основном их комбинации*. А интерпретации различных комбинаций по большей части выводятся *путём умозаключений*, построенных на синтезе знаний относительно природы отдельных шкал. Такая «логическая» стратегия построения интерпретаций имеет два серьёзных недостатка.

Во-первых, полученные таким образом интерпретации далеко не всегда улавливают существо нового содержания, привносимого сочетанием показателей. Некоторая комбинация потому и приобретает дополнительную практическую ценность по сравнению с изолированными показателями, что «снимает» качественно новое содержание. Это содержание часто отражает некие вторичные свойства, которые могут качественно отличаться от исходных составляющих. Поэтому *любые интерпретируемые комбинации показателей должны быть валидизированы как самостоятельные единицы*. Например, корректная трактовка типовых профилей требует отдельной валидации каждого профиля.

Во-вторых, логическая стратегия составления интерпретаций не даёт никаких надёжных оснований для того, чтобы выделить из многочисленных вариаций полезные для практики комбинации показателей. При этом важно напомнить, что разные комбинации неодинаково полезны для разных практических целей. Ценность же разных сочетаний показателей для решения конкретных практических задач может подтвердить только внешняя валидизация. Например, важно эмпирически доказать, что некий профиль действительно характерен для успешного руководителя или менеджера по продажам.

Похожая ситуация складывается и с интерпретацией *производных показателей* (все возможных индексов и коэффициентов). Производные показатели вычисляются на основе исходных (например, на основе оценок по отдельным шкалам) путем некоторых математических операций. Это может быть простое отношение величины одной шкалы к другой или же весьма сложное математическое преобразование. По сути, введение производного показателя – это попытка зафиксировать некоторую полезную комбинацию, придав ей количественное выражение. Такой подход весьма продуктивен, поскольку позволяет не только констатировать наличие некоторого

нового содержания, но и даёт возможность измерять его.

Однако чтобы использовать эту вновь появившуюся возможность, любой производный показатель должен быть не только самостоятельно валидизирован, но и *стандартизован*. Только отдельная стандартизация позволит корректно интерпретировать его величину. В реальности же мы постоянно сталкиваемся с практикой интерпретации количественных градаций таких показателей по типу: «если $k < 1$, то...; если $k > 1$, то...». Возникает вопрос: действительно ли это так? И где та грань, которая могла бы отметить качественные изменения в свойстве, которое отражает коэффициент? Дело в том, что ответить на эти вопросы путём логических умозаключений невозможно...

К сожалению, в большинстве случаев в отечественной практике (и, кстати, не только в отечественной) производные показатели в методиках не валидизируются и не стандартизируются, а их интерпретация производится на основании логических предположений, и, по сути, может иметь лишь статус гипотез.

Если говорить о *шкалах, полученных в результате факторизации*, то здесь ситуация более благоприятная: как правило, большинство таких шкал проводятся через процедуры валидизации и стандартизации. Тем не менее их интерпретация (особенно факторов II порядка) оставляет желать лучшего. По всей вероятности, причина здесь в недостаточно тщательной валидизации. Выше уже говорилось, что обычно фактору сначала приписывается название, а затем уже проводится его валидизация. В большинстве случаев эта процедура заключается в подтверждении нескольких корреляционных взаимосвязей, которые можно с натяжкой привязать к обозначенному в названии содержанию. Достаточно ли этого? Следует напомнить, что именно для идентификации вновь полученных факторов было введено понятие «конструктивная валидность» [1, с. 134], изучение которой предполагает серию серьёзно спланированных исследований, к которым привлекаются разнообразные хорошо продуманные методы. Только такая проработка может обеспечить полноценную и корректную интерпретацию факторных шкал.

Для разъяснения сути измеряемого параметра часто бывает полезно показать его связи с другими явлениями. Для этого в тексты вводят ссылки на *другие переменные, связан-*

ные с диагностируемой. Например, интерпретация результатов теста карьерных ориентаций может сопровождаться утверждениями, подобными следующим: «Люди, ориентированные на предпринимательство, имеют высокую мотивацию достижения...» или «С возрастом и опытом работы ориентация на менеджмент становится сильнее...». Хотя подобные трактовки и кажутся логичными и закономерно вытекающим из содержания конструкта, они должны быть проверены на практике, и, что особенно важно, подтверждены именно для конкретной методики. Например, надо иметь достоверные эмпирические доказательства того, что для людей, у которых выявлена предпринимательская ориентация с помощью методики «Х», также характерна и высокая мотивация достижения.

В современной практике (кстати, не только у нас в России, но и за рубежом) в тексты интерпретаций императивом вводят такие положения, которые на самом деле являются лишь гипотезами и должны проверяться при изучении конструктивной валидности, прежде чем попасть в интерпретации. Особой «виртуозности» в этом смысле достигают интерпретаторы при толковании данных проективных тестов. В описании некоторых комбинаций результатов (например, теста Роршаха или проективного цветового теста) можно насчитать до десяти различных новых свойств, черт, конструктов, которые вводятся для объяснения содержания. Хорошо, если из них валидизированы хотя бы один-два.

Такие свободные трактовки приводят к генерализации интерпретаций – распространению их на содержательные области, которые сам тест не затрагивает. Отсюда иллюзия у новичков и непрофессионалов – что одним тестом можно объять необъятное, решить много разноплановых проблем и описать чуть ли не всю личность человека. Важно, чтобы разработчик мог чётко обозначить верифицированные на настоящий момент границы валидности теста. И именно в пределах этих границ следует выдерживать интерпретации.

В руководстве (или в приложении к нему) полезно приводить конкретные эмпирические свидетельства тех утверждений, которые включены в интерпретации. Это важно и в плане формирования психометрической культуры пользователя тестов. В настоящее время пользователь не имеет средств отличить профессионально разработанный инструмент от кустарного творчества, поскольку не знает, на

что ему ориентироваться, и не приучен искать в руководствах научно обоснованных подтверждений предлагаемым интерпретациям.

Все вышеизложенные принципы построения интерпретаций и отчётов обеспечивают сохранение валидности теста в процессе его последующей эксплуатации. Учитывая эти принципы, рассмотрим конкретные действия разработчика, которые должны быть предприняты для решения ключевых задач «интерпретационного» этапа создания методики. Напомним, что таких задач три:

- 1) разработка схем интерпретации,
- 2) подготовка текстов интерпретаций,
- 3) описание моделей составления отчётов.

1. Разработка схем интерпретации

Схема интерпретации представляет собой *последовательность действий, необходимую для извлечения из полученных результатов сведений, полезных для конкретной практической цели*. Напомним, что для многоцелевых тестов таких схем может быть несколько, и каждая из них должна быть описана отдельно.

Для того чтобы разработать полезную для практики схему интерпретации, необходимо решить ряд промежуточных задач:

- а) построение общей стратегии интерпретации,
- б) определение ключевых показателей и комбинаций,
- в) выстраивание последовательности шагов анализа и интерпретации.

Вопросы, касающиеся *построения общей стратегии интерпретации*, являются ключевыми при разработке будущей схемы. Разные задачи практики (отбор, прогноз, консультирование по развитию и др.) отдают предпочтение разным стратегиям интерпретации. При построении стратегии решаются вопросы, касающиеся выбора основной линии анализа. Например, акцент может быть сделан на сопоставлении результатов с неким критериальным эталоном, или на категоризации результатов (отнесении к группе, типу), или же на многостороннем описании индивидуальной картины [10, 12, 16]. Стратегии интерпретации также различаются в зависимости от того, с какими нормами соотносятся полученные результаты (с групповыми, критериальными, предметно-ориентированными) [4]. Стратегическим является и вопрос выбора единиц интерпретации. Например, для хода анализа немаловажно, будут ли рассматриваться показате-

тели по отдельным шкалам с последующей их «склеивкой» или же единицами интерпретации будут служить устойчивые комбинации, паттерны результатов (такие, как типовые профили и т. д.) [12, 16, 19]. Различаются также и стратегии интерпретации при рассмотрении индивидуальных и групповых результатов. Некоторые задачи практики (например, разработка организационных мероприятий) обычно опираются на интерпретацию групповых результатов диагностики [19].

Напомним, что для тестов, которые планируется использовать для нескольких практических целей, создаётся несколько отличающихся друг от друга схем интерпретации, в основе каждой из которых лежит своя стратегия.

Определение ключевых показателей и комбинаций также производится отдельно для каждой из интерпретационных схем. Эта задача особо актуальна для сложных многошальных методик, а также для тестов, задающих неоднозначную трактовку результатов (например, проективных). Далеко не вся информация, предлагаемая тестом, может быть полезна для конкретной практической цели. Поэтому возникает необходимость выделить из многочисленных вариаций именно *те комбинации показателей, которые несут полезную информацию для решения конкретных практических задач*. Они и станут ключевыми для интерпретации.

Как правило, в рамках одной схемы интерпретации ключевых показателей выделяется относительно немного, но именно они служат ориентирами и источниками информации при принятии диагностических решений.

В качестве ключевых могут выступать весьма разные паттерны результатов, и далеко не всегда это просто высокие или низкие баллы по какой-то шкале. Например, в качестве ключевых могут быть выделены:

- достижение заданного уровня одновременно по двум (или более) шкалам,
- определённая величина отношения одного показателя к другому,
- характерное соотношение двух показателей при заданной величине третьего,
- сглаженность – рельефность профиля или гексограммы,
- особая конфигурация профиля и др.

Как правило, ключевые показатели тесно связаны с эмпирической реальностью и отражают такие особенности результатов, которые имеют «осязаемые» последствия в

жизни и значимы для принятия практических решений.

Опытные пользователи часто самостоятельно приходят к постепенному вычленению таких показателей. Происходит как бы спонтанная внешняя валидизация методики пользователем, поскольку результаты постоянно применяются им на практике. В исследовании К.Р. Червинской подчёркивается, что в основе экспертного умения диагноста лежит многолетняя практика решения проблем [11,12]. У опытного пользователя формируется своя система «диагностических понятий», которая как раз и отражает ключевые показатели. В такую систему могут входить не только паттерны результатов и их сочетания, но и продуктивные способы разбиения шкалы на интервалы, способы определения, является ли показатель пиком в профиле и др. [12]. Опытный диагност знает, на что следует обратить внимание в первую очередь, что важно, что критично в полученной картине результатов. Однако как показали исследования К.Р. Червинской, такой опыт трудно вербализовать и передать другим: требуются специальные процедуры «извлечения экспертных знаний», позволяющие выявить и формализовать продуктивные решения [11,12].

В то же время надо отметить, что использование знаний экспертов чрезвычайно важно при обучении начинающих пользователей. Уже существует опыт создания автоматизированных экспертных систем, построенных на основе моделирования рассуждений экспертов в процессе интерпретации результатов тестирования (К.Р. Червинская). Также пользуется популярностью так называемый «прецедентный» подход (И.Г. Залюбовский, А.Г. Шмелёв), в основе которого лежит накопление базы реальных диагностических случаев («прецедентов»), часто с завязкой на критериальные данные (например, на информацию об эффективности профессиональной деятельности обследованных). Указанные подходы («извлечение экспертных знаний» и «прецедентный»), несмотря на их различие, могут быть использованы для выделения ключевых показателей для конкретных интерпретационных схем. Однако это актуально для уже существующих методик, опыт применения которых достаточно широк и имеет длительную историю. Создание экспертных систем на основе указанных подходов может в высшей степени улучшить пользовательские характеристики таких методик.

Для вновь создаваемых методик задача определения ключевых показателей должна быть возложена в большей мере на разработчиков. По существу, такая задача должна входить в программу психометрической проверки методики. Автору, несомненно, легче указать стратегию поиска продуктивных решений при интерпретации и выделить полезные ключевые показатели, поскольку он знает методику «изнутри», тщательно изучив все её нюансы и детали в процессе разработки. (Это не исключает последующее дополнение найденных решений на этапе эксплуатации методики с участием других исследователей и непосредственных пользователей). При этом особо важно, чтобы все ключевые показатели были валидизированы. Тогда рекомендации автора будут опираться не на интуицию, а на обоснованное знание особенностей методики, на достоверные эмпирические результаты, подтверждающие ценность тех или иных паттернов для решения конкретных практических задач. В большинстве случаев найденные ключевые показатели (и даже сложные комбинации) можно зафиксировать либо с помощью критериальных норм, либо в виде производных коэффициентов с их последующей стандартизацией.

Выделение для каждой интерпретационной схемы ограниченного числа полезных ключевых показателей делает процесс интерпретации для пользователя более конкретным, экономичным и технологичным.

Выстраивание последовательности шагов анализа и интерпретации

Когда ясна общая стратегия интерпретации и выделены ключевые показатели, для окончательного оформления интерпретативной схемы необходимо прописать последовательность действий при анализе и интерпретации результатов. Последовательность действий фиксируется в виде шагов, регламентированных правилами вывода (например, «Если... – то...»). Традиционно такая задача в наиболее явном виде прорабатывалась при создании компьютерных систем интерпретации, где выступала необходимым условием автоматизации процесса. Однако и для методик, не предполагающих компьютерного обеспечения, формализация схем интерпретации является желательной, поскольку снижает общую сложность методики для пользователя [5].

Последовательность шагов зависит от выбранной стратегии, состава ключевых по-

казателей, а также во многом определяется характером исходных данных. В частности, от характера исходных данных зависит сложность модели и применяемые способы структурирования. Примером элементарного уровня сложности может быть линейная схема, построенная для интерпретации одношкального теста. В этом случае действия пользователя заключаются лишь в том, чтобы, ориентируясь на нормы, отнести результат конкретного тестируемого к одному из диапазонов и выбрать соответствующий текст описания из нескольких имеющихся в наличии. Высокую сложность для пользователя представляют многошкальные методики и батареи тестов, где интерпретация требует учёта внутренних взаимосвязей между параметрами и согласования противоречивых данных [5, 11]. Здесь (как уже указывалось выше) особое значение имеет предварительная проработка ключевых показателей, в качестве которых чаще всего выступают паттерны результатов. Схемы интерпретаций для таких методик должны учитывать противоречия в данных, допускать параллельное ведение нескольких линий анализа и вариативность решений. При их построении могут быть использованы сложные модели структурирования (такие, как деревья решений с множественным ветвлением, сети и др.) [12].

Однако важно отметить, что практическая ориентация конкретной интерпретационной схемы частично решает проблему избыточной сложности. Подчеркнём ещё раз, что особенность практико-ориентированных схем интерпретации состоит в том, что их *центральными структурными элементами выступают* не все подряд, а именно *ключевые показатели*, полезные для решения той практической задачи, на которую ориентирована схема интерпретации. Как уже говорилось выше, таких показателей и комбинаций оказывается немного. (Остальные показатели, многие из которых представляют ценность для каких-то других целей, в рамках конкретной схемы опускаются). Поскольку практико-ориентированная схема построена на пошаговом продвижении «в поле» ограниченного числа ключевых показателей, структура отдельной схемы, как правило, не превышает порога умеренной сложности. Это в большинстве случаев позволяет представить схему интерпретации графически, что удобно для пользователя.

Таким образом, конкретная методика может быть обеспечена несколькими самостоя-

тельными схемами интерпретации, предназначенными для разных целей диагностики. В основе каждой из таких схем лежит своя стратегия интерпретации, а узловыми элементами является свой набор ключевых показателей. При этом ограниченное число полезных показателей делает каждую схему достаточно экономичной, а полученные с её помощью интерпретации более конкретными и функциональными в смысле решения поставленных практических задач.

2. Подготовка текстов интерпретаций

По определению, интерпретировать – значит «...истолковывать, объяснять, излагать более понятным языком, раскрывать смысл, содержание чего-либо» [9, с. 245]. В этом смысле интерпретация результатов теста – это перевод данных, в большинстве случаев выраженных в числовом формате, на язык текстов, доступных пользователю, удобных для восприятия и пригодных для принятия практических решений по результатам диагностики.

Подготовка текстов интерпретаций – отдельная задача VI «интерпретационного» этапа, к которой приступают непосредственно после разработки схем интерпретации. Разработчик должен предложить пользователю толкование различных вариантов полученных результатов.

Степень формализации текстов интерпретаций может быть различна. Это могут быть полностью законченные тексты, которые позже будут использованы в отчётах практически без изменений, или же отдельные текстовые фрагменты, из которых «собирается» отчёт. Также могут быть полезными и мало структурированные вспомогательные тексты, разъясняющие суть тех или иных показателей. Ориентируясь на такие тексты, пользователь будет писать отчёты самостоятельно.

Традиционный подход, когда пользователю предлагаются лишь весьма абстрактные описания «высоких» и «низких» баллов по отдельным шкалам, является малопродуктивным и недостаточным для практики принятия решений. Хорошо проработанные тексты интерпретаций должны содержательно «обслуживать» разработанные ранее схемы интерпретации. Для одной и той же методики тексты, сопровождающие разные интерпретативные схемы, могут варьировать по составу и содержанию.

Можно выделить несколько видов текстов, которые рекомендуется помещать в ру-

ководство пользователя. Разные виды текстов имеют разное целевое назначение:

а) *базовые тексты* – это тексты, назначение которых сформировать *общее концептуальное понимание* измеряемого явления и его структуры. К базовым текстам относятся определение центрального конструкта и описание основных структурных элементов (например, отдельных шкал). Для пользователя формулируются исходные понятия и акцент делается на разъяснении заложенного в тест содержания [2, 12]. Если здесь используются трактовки «высоких» и «низких» баллов по шкале, то они больше выполняют иллюстративную роль, способствуя лучшему пониманию сути измеряемого явления. Базовые тексты инвариантны для всех интерпретационных схем внутри одной методики;

б) *служебные тексты* – это тексты с трактовками результатов, *соответствующих ключевым показателям и комбинациям, разработанные для конкретной схемы интерпретации*. Они предназначены для «обслуживания» такой схемы и наполняют описаниями её узловые ячейки. Поэтому такие тексты вариативны, и для одной методики могут быть подготовлены несколько наборов служебных текстов, предназначенных для обеспечения разных интерпретационных схем. Примерами служебных текстов могут быть описания отдельных профилей или сложных комбинаций нескольких показателей, важных для принятия практических решений. Также это могут быть описания характерных профилей, стилей, типов и т.д.

Важно отметить, что, в отличие от базовых, служебные тексты более конкретны, практико-ориентированы и требуют чёткого обозначения формальных границ ключевых показателей. (Например, должно быть указано, что данное описание соответствует случаю, когда «...показатели X и Y одновременно достигают 7 баллов и более при условии, что Z при этом не ниже 4...»). Если в качестве ключевых показателей выступают определённые диапазоны значений шкалы (например, от 2 до 5 или более 8 баллов), то для каждого такого диапазона составляются свои отдельные тексты [12]. Как уже подчёркивалось выше, все ключевые показатели должны быть предварительно валидизированы: в этом случае сложностей с их корректным описанием не будет.

Именно служебные тексты являются непосредственной основой для составления ди-

агностических отчётов. Именно из них будет собираться итоговая картина тестовых результатов. Поэтому служебные тексты должны быть проработаны особо тщательно и, по возможности, формализованы. Это особенно важно, если планируется создание автоматизированной системы интерпретации и составления диагностических отчётов;

в) *дополнительные тексты содержат информацию, полезную для лучшего понимания природы измеряемого явления или для принятия практических решений*. Это могут быть тексты, в которых описываются переменные, взаимосвязанные с измеряемыми, или же обозначаются факторы, вызывающие изменения в показателях [19,20]. Например, в качестве дополнительных, в руководство пользователя могут быть включены тексты, касающиеся сведений о гендерных или возрастных различиях, или же об источниках возможных искажений в тестовых результатах. Также в дополнительных текстах описываются переменные–модераторы, усиливающие или сглаживающие критериальный эффект и влияющие на прогноз. В раздел дополнительных текстов могут быть помещены нормативные данные. Полезно, если в дополнительных текстах приводятся эмпирические свидетельства утверждений, положенных в основу тех или иных интерпретаций.

Дополнительные тексты составляются в свободной форме, удобной для восприятия. Пользователь может обращаться к ним избирательно в зависимости от требований ситуации. При этом некоторые фрагменты таких текстов могут быть использованы и в диагностических отчётах. Например, хорошо, если заказчик, применяющий результаты диагностики для целей отбора, будет осведомлён о переменных–модераторах, которые могут повлиять на прогнозируемый исход и на вероятность достижения критериального поведения конкретным претендентом на должность.

Надо отметить, что количество дополнительных текстов, сопровождающих методику, может расти с опытом её использования и накоплением дополнительных данных, проливающих свет на особенности конструкта.

Таким образом, на VI «интерпретационном» этапе создания методики разработчиком должны быть подготовлены три вида текстов, различающихся своим назначением и степенью формализации. Некоторые из них будут непосредственно использованы при составлении диагностических отчётов, другие служат

вспомогательным целям, способствуя лучшему пониманию картины полученных результатов и их корректной интерпретации.

3. Описание моделей составления отчётов

Завершающая задача VI «Интерпретационного» этапа заключается в разработке и описании моделей составления диагностических отчётов.

Диагностический отчёт представляет собой текстовый (или электронный) документ, предназначенный для конкретной категории лиц. В этом документе представлена ключевая информация, извлечённая из результатов диагностики и необходимая указанной группе лиц для решения конкретных практических задач. Диагностический отчёт составляется на основе выполненной интерпретации конкретных результатов и с использованием подготовленных разработчиком текстов.

Традиционно ответственность за составление отчётов возлагалась на пользователя. Отсутствие каких-либо внятных рекомендаций по составлению диагностических отчётов превращало задачу составления отчёта, скорее, в искусство или же сложную «игру» с заказчиком. При необходимости составить неформализованные характеристики, прогнозы или дать комплексное заключение и рекомендации по отбору и развитию пользователей оказывался в весьма сложной ситуации и мог рассчитывать только на свой опыт, талант и практическую хватку.

Однако в последнее время важность этой части работы всё больше осознаётся психологическим сообществом и ответственность за качество диагностических отчётов, сопровождающих тест, постепенно возлагается на плечи разработчиков [13, 15, 17]. В частности, к отчётам так же, как и к самому тесту стали предъявляться требования серьёзной эмпирической проверки [13, 15, 19]. Валидность диагностического отчёта стала одним из важных компонентов при оценке качества диагностической методики (см. форму рецензии на тест EFPA, раздел 4). В стандарты AERA, APA & NCME включено немало пунктов, посвященных именно диагностическим отчётам [13].

Диагностический отчёт становится итогом, зримым результатом всей работы, связанной с созданием методики. По сути, именно по качеству отчёта заказчик делает выводы о качестве диагностического инструмента. Поэтому тест, даже самым тщательным обра-

зом разработанный, но не снабжённый современными средствами представления результатов, не завоеует должного признания и не будет востребован.

Главное назначение отчёта – корректно донести информацию о результатах тестирования до заинтересованной в этом группы лиц. При этом отчёт должен быть, с одной стороны, информативным и полезным для практики, с другой – лёгким для восприятия. Поэтому *способ подачи результатов* становится одной из важных составляющих качества диагностического отчёта.

В последние годы за рубежом появились многочисленные исследования, направленные на изучение восприятия и понимания отчётов разными категориями пользователей [17, 19, 20, 21, 22]. Обобщая полученные в этих исследованиях результаты, можно выделить основные *качества эффективных отчётов*.

В первую очередь выделяются *вопросы организации информации* в отчётах:

- в отчётах не должно содержаться слишком много данных (следует приводить только основные, важные для практики);

- данные должны быть сгруппированы эффективным, несущим полезную информацию способом;

- разные средства наглядности (графики, рисунки) следует использовать обдуманно, преимущественно для того, чтобы подчеркнуть наиболее важные результаты;

- желательно несколько вложенных друг в друга уровней отчёта (различающихся по глубине и подробности изложения информации), верхним уровнем должно быть «summary», где результаты представлены в наиболее общем виде.

Также пользователи считают, что отчёт должен содержать информацию, позволяющую им самостоятельно оценить результаты, полученные конкретным тестируемым, увидеть его *уровень на фоне других*. В качестве такой информации называют:

- нормативные данные (нормативные интервалы, текущие баллы и т. д.);

- данные для сравнения (например, таблицы с указанием процента лиц из разных социальных или профессиональных групп, выполнивших тест на определённом уровне);

- описание ожидаемых (критериальных) результатов;

- описание недостатков, ошибок, отклонений в показателях конкретного тестируемого.

Важным считается включение в отчёты разделов, *разъясняющих содержательные вопросы*. Это:

- сообщение цели тестирования;
- описание основных параметров измерения (центрального конструкта и шкал или субтестов);

- сопровождение отчёта словарём терминов.

Также желательна *информация о возможных ошибках в выводах*. Это может быть:

- информация об источниках искажений тестовых результатов;

- дополнительные тексты с альтернативным объяснением полученных результатов.

Среди качеств, повышающих *практическую применимость* отчёта, чаще всего называют следующие:

- отчёт должен «обслуживать» конкретные потребности той группы лиц, для которой он разработан (администраторы школ, родители, педагоги, работодатели и т.д.);

- отчёт должен содержать подробные рекомендации, касающиеся действий, основанных на полученных результатах;

- желательно включение в отчёт дополнительной информации, помогающей принятию решений (например, о переменных-модераторах, влияющих на прогнозируемый исход).

Немаловажными считаются и качества отчёта, обеспечивающие *лёгкость его восприятия и привлекательность*:

- стиль и язык отчёта должен соответствовать возможностям той группы лиц, для которой он предназначен;

- нежелательно использование специальной терминологии (особенно, статистической);

- следует уделять внимание размеру шрифта и использованию физического пространства для размещения текста;

- наглядное представление результатов должно помогать пониманию (например, цветовые решения будут полезны, если эргономичны и используются в умеренном количестве);

- приветствуется визуальная привлекательность оформления;

- общий объём отчёта не должен быть слишком большим.

Перечисленные выше качества отчётов делают их более эффективными как средство сообщения результатов, поэтому должны учитываться при разработке моделей отчётов.

Модель диагностического отчёта – это некоторая фиксированная форма, которая наполняется конкретным содержанием в зави-

симости от полученных результатов. «Каркас» модели составляют её структура (разделы и их последовательность) и способ организации фактического материала. Также особенности модели определяются её стилем, языком и формой, в которой предоставляются практические рекомендации. Все эти характеристики модели определяет и фиксирует разработчик.

В любой модели всегда есть инвариантные элементы, которые не зависят от специфики полученных результатов. Это базовые тексты с описанием основных содержательных единиц, словарь терминов, информация о нормативных данных, шаблоны таблиц, диаграмм, рисунков, тексты с дополнительной информацией, помогающей принятию решений. Все эти элементы отчёта могут быть заранее подготовлены разработчиком, и будут включены пользователем в итоговый отчёт практически без изменений.

В настоящее время не существует законченной классификации моделей диагностических отчётов. Но эта задача представляется весьма актуальной и её решение позволило бы во многом оптимизировать труд разработчиков психодиагностического инструментария на интерпретационном этапе: можно было бы уже говорить о *выборе* моделей, а не об их разработке в каждом конкретном случае. Однако создание такой классификации требует времени и совместной работы специалистов разных профилей. Тем не менее уже можно сформулировать некоторые *общие рекомендации по составлению моделей диагностических отчётов*.

Как указывалось выше, конкретная методика может сопровождаться несколькими схемами интерпретации, предназначенными для разных практических целей. Важно, что для каждой схемы разрабатываются свои модели отчётов. По форме это могут быть характеристики, рекомендации, прогнозы, аналитические справки и т.д. [6, 19]. В первую очередь решается *вопрос определения круга лиц, которым будут сообщаться результаты тестирования*, – от этого зависит количество необходимых моделей. (В большинстве случаев для каждой группы лиц предполагается своя форма отчёта).

Адресное назначение отчёта является ключевым параметром при выборе структуры, состава материала и способа подачи. Прежде всего, здесь следует различать две категории лиц, получающих диагностические отчёты.

Первая – те, кто непосредственно будет использовать результаты тестирования для решения каких-либо практических задач. Для этой категории лиц составляются *основные отчёты*, наиболее подробные, развёрнутые, с большим весом раздела практических рекомендаций. Вторая категория лиц – это люди, которые будут просто осведомлены о результатах. В большинстве случаев, это сам тестируемый и иногда третьи лица (например, родители, руководитель или др.). Для них составляют *ознакомительные отчёты* – более краткие, где акцент делается на обслуживании специфических интересов этих лиц.

Например, в случае диагностики профессиональных компетенций с целью отбора кандидатов на должность основной отчёт составляется для сотрудника отдела кадров, а ознакомительные отчёты – для самого тестируемого и руководителя организации. При использовании теста учебных достижений в школе основными заказчиками могут быть школьный психолог и учитель, а лицами, которые заинтересованы в том, чтобы получить ознакомительную информацию – сами ученики, родители и администрация школы.

Для каждой адресной группы предполагается своя модель отчёта. Даже ознакомительные отчёты, несмотря на свою краткость, могут серьёзно различаться по форме и содержанию. Например, для учеников они пишутся «лёгким» языком, в наиболее общем виде и в позитивном ключе, для родителей – в форме краткого информирования и могут включать рекомендации по развитию ребёнка, для администрации полезны аналитические справки и сводки по классам, параллелям, предметам и т. д.

Любой диагностический отчёт обязательно должен включать *описание результатов тестирования*. Форма представления результатов закладывается в модель. При этом выбор адекватной, информативной и удобной для восприятия формы – весьма непростая задача. С одной стороны, выбор формы определяется характером исходных данных и целью тестирования, с другой – способ подачи должен соответствовать возможностям той группы лиц, для которой предназначен отчёт. Поэтому рекомендуется предварительная пилотажная проверка моделей отчётов для разных категорий пользователей [17, 19]. Ценную информацию для выбора формы представления результатов могут дать фокус-групповые исследования с участием потенциальных пользователей и

клиентов. J.M. Ryan приводит подробное описание такого исследования (The south carolina/serve study of score reporting formats), где изучались шесть разных форматов представления результатов тестов достижений [19]. К исследованию привлекались группы учителей, администраторов школ, специалистов по ассессменту, родителей и др. Респондентов просили оценить предлагаемые форматы с помощью специально разработанной шкалы и стимулировали содержательные дискуссии. Обсуждение базировалось на вопросах о том, чем каждый формат может быть полезен; как конкретно можно использовать предлагаемую в таком виде информацию; какие проблемы могут возникнуть у пользователя, ориентирующегося на тот или иной формат; как можно улучшить способ представления результатов и т. д. [19].

Важной частью отчёта являются *практические рекомендации*. Они наиболее вариативны, и их содержание определяется целями и интересами каждой группы лиц, получающих отчёт. Например, при использовании тестов достижений в школе психолог и учитель решают разные практические задачи, соответственно нуждаются и в разных рекомендациях. В частности, для психолога могут быть полезны рекомендации по развитию учеников, для учителя – по организации учебного процесса. Образцы практических рекомендаций для каждой модели отчёта также должны быть подготовлены разработчиком. Полезно, если в руководстве приводятся примеры отчётов с рекомендациями, составленными для конкретных случаев из практики.

Таким образом, задачу описания моделей диагностических отчётов можно считать выполненной, если в рамках каждой интерпретационной схемы разработчик предлагает пользователю несколько моделей, предназначенных для разных адресных групп. При этом каждая модель имеет фиксированную структуру, шаблоны оформления, свой набор текстовых заготовок, образцы составления рекомендаций, а также комментарии, касающиеся стиля и языка. В зависимости от полученных результатов пользователь наполняет графические шаблоны и таблицы фактической информацией, компоует текст описания результатов из предоставленных текстов (см. выше «Подготовка текстов интерпретаций») и составляет рекомендации с учётом конкретного запроса клиента.

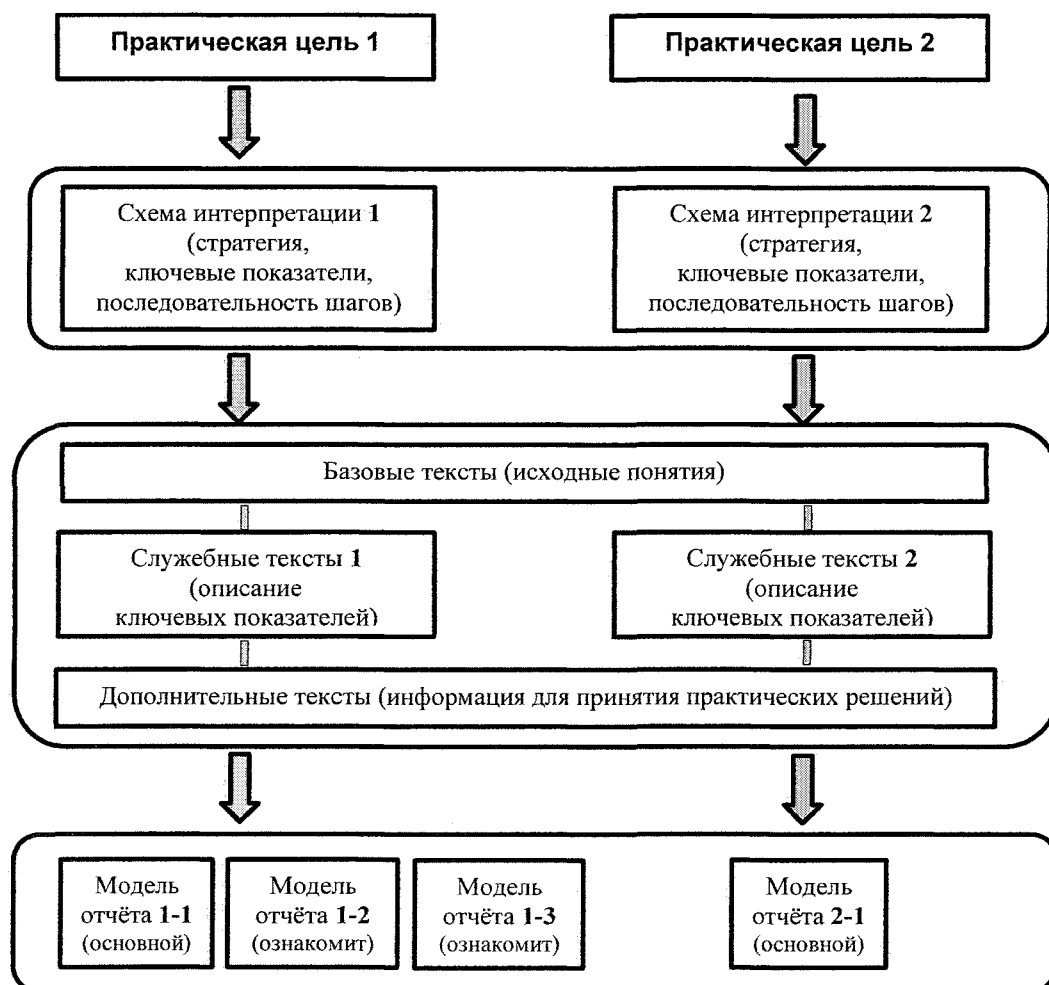
Подводя итоги, можно обобщить изло-

женную выше информацию с помощью условной схемы, отражающей общую структуру системы интерпретации. На рисунке приведена модель системы интерпретации для методики, которую планируется использовать для 2 практических целей. Система интерпретации состоит из 3 блоков, соответствующих 3 описанным выше задачам: 1) схемы интерпретации, 2) тексты, 3) модели отчётов.

Например, такая модель может быть использована для разработки системы интерпретации методики диагностики трудовой мотивации [8]. Первое целевое назначение методики может быть связано с отбором кандидатов на должность, второе – с консультацией самого обследуемого на предмет планирования карьеры. В этом случае предполагается разработка 2 относительно независимых схем интерпретации, базирующихся на разных стратегиях и имеющих в качестве узловых элементов разные ключевые показатели. Так, при отборе на вакантную должность ключевыми будут пока-

затели, полученные в результате соотношения данных тестирования с критериальным профилем, а в случае консультирования по планированию карьеры – особенности ведущих и игнорируемых мотивов, устойчивость полученной иерархии и общая согласованность/противоречивость системы мотивации. Каждая из 2 схем интерпретации обеспечивается своим набором служебных текстов. При этом базовые тексты (для приведённого примера – это общее описание диагностируемых мотивов), а также дополнительные тексты будут инвариантны для обеих схем.

Далее каждая из схем интерпретации сопровождается разработанными специально для неё моделями отчётов. Так, для схемы, соответствующей цели отбора, предполагается 3 отчёта: основной – для сотрудника отдела кадров и два ознакомительных – для руководителя организации и тестируемого. Для схемы, разработанной под задачу консультирования клиента – всего один отчёт – основной, для



Пример общей структуры системы интерпретации для методики, имеющей два целевых назначения

самого человека, прошедшего тестирование.

Если предполагается расширить сферу применения методики (например, использовать её для планирования организационных мероприятий по мотивированию сотрудников), то в систему интерпретации необходимо включить третью схему интерпретации со своей стратегией, ключевыми показателями и моделями отчётов.

В заключение отметим, что предлагаемые пользователю системы интерпретации могут быть формализованы в разной степени. Как уже упоминалось выше, в нашей стране разработаны и апробированы технологии создания полностью автоматизированных экспертных диагностических систем, позволяющих производить интерпретацию и генерировать компьютерные отчёты даже для сложных многошкальных методик (например, К.Р. Червинская, И.Г. Залобовский, А.Г. Шмелёв). Наличие компьютерного отчёта выступает одним из параметров оценки диагностического инструмента (см. форму рецензии ЕФРА). Однако следует понимать, что автоматизация интерпретационных систем облегчает работу для пользователя, но существенно усложняет – для разработчика, перед которым наиболее остро встаёт вопрос проверки каждого шага предлагаемой схемы. Затраты на такую работу могут быть очень велики. Поэтому решение о необходимости и возможности компьютеризации интерпретационной системы должно приниматься самими разработчиком. Возможно, в некоторых случаях (например, при разработке сложных многошкальных методик, проективных тестов, ассессмент-процедур) оптимальным будет выбор частичной автоматизации, когда машина генерирует лишь представление полученных результатов и в диалоговом режиме предлагает варианты трактовок, доступные для внесения корректив, а практические рекомендации составляются уже непосредственно пользователем с учётом нюансов конкретного запроса.

При этом, как уже неоднократно подчёркивалось в статье, любая формализация считается желательной, поскольку снижает общую сложность методики для пользователя [5]. Даже если разработчик решает полностью отказаться от компьютеризации интерпретаций, то всё равно пользователь должен быть снабжён подробным руководством, позволяющим ему пошагово продвигаться, совершая меньше ненужных и ошибочных действий. Это уменьшит количество искажений,

которые могут быть привнесены пользователем в процессе эксплуатации методики, и будет способствовать повышению надёжности и валидности диагностических отчётов, являющихся закономерным итогом всей трудоёмкой работы по созданию методики.

Литература

1. Анастаси, А. Психологическое тестирование / А. Анастаси, С. Урбина. – СПб.: Питер, 2001. – 668 с.
2. Батурин, Н.А. Технология разработки тестов: часть I / Н.А. Батурин, Н.Н. Мельникова // Вестник ЮУрГУ. Серия «Психология». – 2009. – Вып. 6. – С. 4–14.
3. Батурин, Н.А. Технология разработки тестов / Н.А. Батурин, Н.Н. Мельникова // Вестник ЮУрГУ. Серия «Психология». – 2009. – Вып. 7. – С. 11–25.
4. Батурин, Н.А. Технология разработки тестов / Н.А. Батурин, Н.Н. Мельникова // Вестник ЮУрГУ. Серия «Психология». – 2010. – Вып. 8. – Ч. III. – С. 4–18.
5. Батурин, Н.А. Принципы определения сложности психодиагностических методик / Н.А. Батурин, Н.Н. Мельникова // Вестник ЮУрГУ. Серия «Психология». – 2010. – Вып. 8. – С. 36–43.
6. Батурин, Н.А. Новые технологии создания моделей интерпретаций и диагностических отчетов / Н.А. Батурин, Н.Н. Мельникова // Вестник ЮУрГУ. Серия «Психология». – 2010. – Вып. 10. – С. 4–6.
7. Купер, К. Индивидуальные различия / К. Купер; под ред. И.В. Равич-Щербо. – М.: Аспект Пресс, 2000. – 527 с.
8. Мельникова, Н.Н. Диагностика трудовой мотивации / Н.Н. Мельникова, Д.М. Полев // Современная психодиагностика в изменяющейся России: сб. тез. Всерос. науч. конф. / отв. ред. Н.А. Батурин. – Челябинск: Изд-во ЮУрГУ, 2008. – С. 79–81.
9. Ожегов, С.И. Толковый словарь русского языка / С.И. Ожегов, Н.Ю. Шведова. – М.: АЗЪ, 1996. – 928 с.
10. Фер, Р.М. Психометрика: Введение / Фер Р.М., В.Р. Бакарк В.Р.; пер с англ.; под ред. Н.А. Батурина, Е.В. Эйдмана. – Челябинск: ЮУрГУ, 2010. – 445 с.
11. Червинская, К.Р. Психологическая концепция извлечения экспертных знаний на моделях медицинской психодиагностики / К.Р. Червинская // Вестник ЮУрГУ. Серия «Психология». – 2008. – Вып. 2. – С. 68–80.
12. Червинская, К.Р. Психология извлече-

ния экспертных знаний субъектов труда: дис. ... д-ра психол. наук / К.Р. Червинская. – СПб., 2010. – 437 с.

13. American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association. – 101 p.

14. Cizek, G.J. *Standard Setting* / G.J. Cizek // *Handbook of test development*/edited by Steven M. Downing, Thomas M. Haladyna. – 2006, by Lawrence Associates. – P. 225–258.

15. Downing, S.M. *Twelve steps for effective test development* / S.M. Downing // *Handbook of test development* / edited by Steven M. Downing, Thomas M. Haladyna. – 2006, by Lawrence Associates. – P. 3–25.

16. Furr, R.M. *A framework for profile similarity: integrating, similarity, normativeness, and distinctiveness* / R.M. Furr // *Journal of Personality*, October 2008. – P. 1267–1316.

17. Goodman, D.P. *Student test score reports and interpretive guides: Review of current practices and suggestions for future research* /

D.P. Goodman, R.K. Hambleton // *Applied measurement in Education*. – 2004. – 17(2). – P. 145–220.

18. CPI 260 *Coaching report for leaders user's guide* 2002, 2005 by CPP. – 38 p.

19. Ryan, J.M. *Practices, Issues, and Trends in Student Test Score Reporting* / J.M. Ryan // *Handbook of test development*/edited by Steven M. Downing, Thomas M. Haladyna. – 2006, by Lawrence Associates. – P. 677–710.

20. Ysseldyke, J. *Reporting results of student performance on large-scale assessment* / J. Ysseldyke, J.R. Nelson // *Large-scale assessment programs for all students* / edited by G. Tindal & T. M. Haladyna. – 2002, by Lawrence Erlbaum Associates. – P. 467–480.

21. Zwick, R. *An investigation of alternative methods for item mapping in the National Assessment of Educational Progress* / R. Zwick, D. Senturk, J. Wang, S.C. Loomis // *Educational Measurement: Issues and Practice*. – 2001. – 20(2). – P. 15–25.

22. Wainer, H. *Understanding graphs and tables* / H. Wainer // *Educational Researcher*. – 1992. – 21(1). – P. 14–23.

Поступила в редакцию 16 сентября 2010 г.

Батурин Николай Алексеевич. Доктор психологических наук, профессор, декан факультета психологии, Южно-Уральский государственный университет, г. Челябинск: nikbat@list.ru.

Nikolay A. Baturin. PsyD, professor, the dean of the Faculty of psychology, head of chair «Psychological diagnostics and Counseling» of South Ural State University: nikbat@list.ru.

Мельникова Наталья Николаевна. Кандидат психологических наук, доцент кафедры социальной психологии, Южно-Уральский государственный университет, г. Челябинск. Область научных интересов: психология организации жизни, социально-психологическая адаптация личности, системный подход в психологии: MNN17@yandex.ru.

Natalia N. Melnikova. Candidate of Psychological sciences, docent of department of social psychology of South Ural State University. Area of scientific interests: psychology of the organisation of a life, socially-psychological adaptation of the person, the system approach in psychology: MNN17@yandex.ru.