



ПРОБЛЕМЫ КАЧЕСТВА ТЕСТОВЫХ МАТЕРИАЛОВ

А.Н. Величко, О.А. Ефремова

В статье раскрываются подходы к исследованию качества тестовых материалов, определяется значение экспертизы диагностических материалов, представляются проблемы, проявившиеся при практическом экспертировании банков тестовых заданий.

Ключевые слова и словосочетания: диагностика, экспертиза, задания тестового характера, тесты.

Современная педагогическая диагностика в целом, как и диагностика предметной обученности, немыслима без использования тестов или хотя бы комплектов заданий тестового характера. При применении тестовых материалов становятся актуальными следующие проблемы:

- ✓ зачем использовать тесты, каковы место и роль тестовых материалов в педагогической диагностике;
- ✓ как определить параметры, проверяемые тестовыми материалами;
- ✓ как создать качественные тесты;
- ✓ насколько правомерно использовать собственные тесты, не прошедшие сертификацию;
- ✓ как проверить качество имеющихся тестовых материалов;
- ✓ какие выводы правомерно делать при анализе результатов тестирования и т.д.

Однако в данной статье не будем претендовать на разрешение всех проблем, остановимся на вопросе проверки качества тестовых материалов. Качественными считают те тесты, которые прошли процедуру сертификации. Заметим, что сертифицирование материалов может происходить на разных уровнях; самый высокий уровень, конечно, уровень Госстандарта, только такие тесты по формальным показателям являются качественными. Однако практика показывает,

что не всегда готовые тесты устраивают по всем параметрам. Иногда их просто еще нет.

При проведении процедуры сертификации на любом уровне обязательно формируется заключение по тестовым материалам, состоящее из двух элементов:

- ✓ экспертизы, то есть исследования специалистом (экспертом), имеющим специальные познания в исследуемой области знаний (тестологии) тестовых материалов;
- ✓ апробации, которая заключается в проведении пробной процедуры тестирования для определения статистических характеристик теста.

Наличие одного элемента (либо экспертизы, либо апробации), как показывает практика, недостаточно для определения качества используемых тестовых материалов. Интенсивное накопление материалов тестового характера делает проблему оценки их качества весьма актуальной. Особую значимость она приобрела для преподавателей Новосибирского государственного педагогического университета в связи с необходимостью подготовки диагностических материалов для проверки остаточных знаний в рамках проблемы оценки качества подготовки специалиста.

Для упрощения процедуры создания и использования тестовых материалов стихийно и централизованно создаются различные компьютерные программы – тестеры. Каждый пользователь может найти для себя наиболее удобный вариант. Накануне аттестации вуза Новосибирский государственный педагогический университет (НГПУ) выбрал для внедрения систему тестирования, используемую в процедуре внешнего аттестационного тестирования, разработанную Московским Центром тестирования профессионального образования (модель АСТ-Центра). Внедрение системы в полном объеме предполагает интенсивный процесс создания и включения в работу со студентами диагностических материалов в тестовой форме: наполнение, так называемых, банков тестовых заданий (БТЗ); создание на их основе диагностических тестов с различными целями проверки. Хотя материалы, полученные при взаимодействии с АСТ-Центром, объемны и многоплановы, в них достаточно много места отведено описанию структуры экспертизы, практика вносит свои коррективы и ставит новые проблемы.

Очевидно, что экспертиза, проводимая недостаточно качественно и без привлечения соответствующих специалистов, а уж тем более её отсутствие, приводят к тому, что при работе с БТЗ и составлением тестовых заданий накапливается значительное количество диагностических проблем, уменьшается качество и увеличиваются последующие трудозатраты, связанные с исправлением недостатков. Соответственно, проведение экспертизы не может быть формальной процедурой, и она должна быть обязательно многоуровневой. Выделяют следующие уровни:

- ✓ авторская экспертиза, или самоэкспертиза, которая обязательно проходит в процессе разработки всего БТЗ и составления отдельных тестовых заданий. Она представляет собой действия по самоконтролю и самооценке, устраняющие неизбежные, часто произвольные ошибки разработчика тестовых материалов. После завершения работы над созданием БТЗ проводится его экспертиза для выявления соответствия содержанию требований, предъявляемых к БТЗ;

- ✓ внутренняя экспертиза, проводимая внутри образовательного учреждения специалистами, не принимавшими участия в разработке данного БТЗ, но являющимися специалистами в содержательной области;

- ✓ внешняя экспертиза, проводимая специалистами в данной области знаний, которые не являются работниками данного учреждения;

- ✓ внешняя экспертиза в рамках процедуры сертифицирования тестовых материалов, проводимая специалистами в данной области знаний, имеющими права на проведение таких процедур (сертификат эксперта или удостоверение по прохождению курсов). Для получения внешнего экспертного заключения (рецензии) такого уровня можно обратиться в соответствующие УМО.

Каждый уровень экспертизы необходимо оформлять в виде протокола или рецензии, желательно их совмещать. Формализованность протокола с использованием обобщающих таблиц позволяет быстро сравнивать оценки одинаковых позиций у разных экспертов и формировать общее заключение. Кроме того, специальная форма протокола заставляет эксперта не пропускать ни одной позиции при оценке тестовых материалов. Рецензия необходима для пояснения

экспертом собственного впечатления и позволяет описать замеченные неточности, а не только констатировать их наличие. Может возникнуть мнение об избыточности заполнения протокола при проведении авторской экспертизы. Наш опыт позволяет с полной уверенностью говорить о необходимости заполнения этого документа. Во-первых, это позволяет сравнить собственное мнение с мнением другого эксперта. Если знакомишься с заключением эксперта, не проводя самообследование, то мнение другого эксперта дает «наводки» и приводит к искажению собственного мнения. Можно еще поспорить, кто более прав, от своей позиции не следует быстро отказываться. Во-вторых, заполнение протокола позволяет перевести интуитивные ощущения о качестве созданных материалов на осознанный уровень. Собственный опыт показывает, что первые впечатления иногда обманчивы, только методичное исследование качества по определенным параметрам в рамках заданной определенной формы, чтобы не задумываться над систематизацией своих впечатлений, позволяет полноценно оценить качество подготовленных материалов.

Экспертиза неизбежно несет некоторый субъективный характер оценки, поэтому от качества подбора экспертов-предметников и их квалификации во многом зависит успех проведения экспертизы. Очевидно, чем больше экспертная группа, тем более верные результаты оценок могут быть получены. Однако привлечение слишком большого числа экспертов, особенно при отсутствии возможности их отбора, нежелательно. В этом случае мнение каждого специалиста слабо влияет на конечные результаты. Кроме того, оценки компетентных экспертов размываются во множестве оценок малокомпетентных специалистов. Чем больше экспертная группа, тем больше времени затрачивается не только на подбор специалистов, но и на сведение воедино результатов экспертиз и формирование окончательного заключения, что приводит к удорожанию процедуры проведения экспертизы.

Уже появляются специальные разработки по оценке согласованности экспертов, что подчеркивает общий характер проблемы. По мнению сотрудников АСТ-Центра, группа специалистов должна включать от 5-ти до 13-ти человек. На наш взгляд, для проведения внутренней экспертизы достаточно подготовленных

2 – 3-х специалистов, если они не влияют на мнения друг друга, выступают независимыми.

В идеологии адаптивной системы тестирования (АСТ) при проведении экспертизы следует четко различать экспертизы:

- ✓ каждого тестового задания;
- ✓ всего банка тестовых заданий (БТЗ);
- ✓ тестов, составленных на основе БТЗ.

Для полноценной экспертизы необходимо определить ее критерии и выделить параметры, по которым она проводится.

Вполне возможно совместить в одном протоколе экспертизу каждого тестового задания и всего банка тестовых заданий в целом. Для оценки теста необходим другой протокол, так как тест – это не простая совокупность, пусть и структурированная по содержанию, как в БТЗ, тестовых заданий. Тест имеет самостоятельные параметры.

Экспертиза всего банка тестовых заданий обязательно предполагает, что анализируется каждое тестовое задание. Цель проведения экспертизы – проверка тестового задания на соответствие методологическим требованиям. Показателей каждого тестового задания много, не следует игнорировать ни один из них, последовательно оценивая каждое тестовое задание по всем параметрам. По некоторым показателям отсутствуют однозначные критерии, что приводит к субъективизму оценки. Именно поэтому при проведении экспертизы желательно приглашать не менее 2 – 3-х экспертов.

В протоколе экспертизы тестовых заданий, следуя материалам АСТ-центра, предлагается определять несколько параметров:

1. *Уникальный идентификатор тестового задания.* Он необходим для того, чтобы можно было точно отличить и выделить конкретное тестовое задание из всего массива БТЗ.

2. *Код принадлежности тестового задания к дидактической единице* – разделу спецификации БТЗ, которая в своей основе содержит дидактические единицы образовательного стандарта. В этом пункте мнение эксперта, как показала практика, может расходиться с мнением авторов. Для проведения независимого анализа эксперт не должен знать, к какой дидактической единице из спецификации отнесли тестовое задание авторы БТЗ. Этот па-

раметр достаточно информативен, он позволяет исключить неоднозначные задания и задания, проверяющие сразу много дидактических единиц. Исходя из теоретических положений тестологии, одно задание должно проверять одну дидактическую единицу, только в этом случае можно сделать локальную оценку усвоения именно этой дидактической единицы.

3. *Соответствие ТЗ содержанию ГОС* или учебной программе. Даже небольшой опыт компоновки БТЗ показал, что полное соответствие дидактических единиц, определяемых в спецификации тестовых материалов, государственному образовательному стандарту невозможно. Автор БТЗ обязательно корректирует, разворачивает словосочетания, предложения, заявленные в образовательном стандарте. Кроме этого, оказывается, несмотря на единый образовательный стандарт, у экспертов есть собственное мнение по поводу его трактовки.

4. *Уровень значимости содержания* тестового задания. Эксперт на свое усмотрение определяет данное задание: сомнительное, приемлемое (допустимое), важное или существенное по содержанию.

5. *Выполнение требований к составлению тестовых заданий*. Наблюдая за работой экспертов, можно сделать заключение, что этот параметр могут оценить только специально подготовленные эксперты. Прописывание всех требований к составлению тестового задания в пояснительной записке к протоколу занимает очень много места и времени.

6. *Максимально допустимое время предъявления задания* на экране дисплея, в минутах. Нами этот параметр был модифицирован в оценку времени, которое требуется среднему испытуемому для ответа на задание. Оказалось, что по этому параметру, с одной стороны, мнения экспертов значительно расходятся, с другой стороны, преподаватели, учителя, выступающие экспертами, часто неадекватно его оценивают, тем самым необоснованно увеличивая темп прохождения теста.

7. *Форма ТЗ* – открытая, закрытая, на упорядочение, на соответствие. В этом пункте расхождений не бывает, формы заданий опознаются однозначно, но анализ этого параметра «по вертикали» – от задания к заданию, позволяет сделать заключение о сбалансиро-

ванности разнообразия форм. При однообразной форме тестовых заданий этот пункт из экспертного заключения можно исключить.

8. *Количество концептов* (логических шагов или интеллектуальных действий), необходимых для решения тестового задания. Если задание имеет несколько вариантов решения, традиционные и нетрадиционные подходы к решению, то при внесении в протокол количества концептов следует остановиться на наиболее стандартном варианте решения, опирающемся на основные (фундаментальные) законы или понятия, имеющиеся в государственном образовательном стандарте. Очевидно, что даже в сложном тестовом задании количество концептов не должно превышать трех. Практика показала, что оценка этого параметра затруднительна и субъективна. Получается, чем выше компетентность специалиста в данной области знания, тем меньше концептов он ставит. Очевидно, у самого эксперта знания и действия достаточно свернуты, он оперирует не исходными законами, а привычными штампами, что приводит к сворачиванию логических шагов при решении.

9. *Количество дистракторов* в тестовом задании. Для закрытой формы и на упорядочивание указывается количество вариантов ответа в задании. Для заданий на соответствие отмечается количество пар и количество вариантов ответа, оставшихся без пары. Для заданий открытой формы указывается количество возможных правильных ответов, определенных автором, либо вообще не отмечается данный показатель. Нами в практике работы это положение часто игнорировалось.

10. *Глубина иерархии* в спецификации теста, т.е., к какому узлу спецификации данное задание относится: к главе, разделу, подразделу, теме или понятию. По нашему мнению, правильно, с учетом иерархии, приписанный код тестового задания позволяет игнорировать этот показатель.

11. *Форма мышления* (суждения) определяет уровень владения проверяемым данным тестовым заданием знанием:

- ✓ знание знакомства (опознание);
- ✓ знание копии (воспроизведение);
- ✓ знания умения (применения).

Этот параметр можно интерпретировать как определение уровня деятельности. Возможно, целесообразнее более конкретно

указывать вид проверяемого умения. Внесение в протокол экспертного заключения именно проверяемого умения показало, что авторы БТЗ неосознанно отдают предпочтение проверке лишь одного умения, чаще всего, репродуктивного характера. Очевидно, опираясь на тезис: чему научил, то и проверяю.

12. *Мера трудности*, качественная оценка, представляет собой субъективный показатель, как и само понятие «трудность». При оценке не следует смешивать трудность и сложность, ведь последнюю косвенно задает количество концептов. Сложность задания можно количественно определить по количеству логических шагов, используемых при ответе на задание. В протоколе представлено субъективное мнение эксперта – задание легкое, средней трудности, трудное.

13. *Однозначность* ТЗ предполагает исключение всякого разночтения в формулировке как самого задания, так и ответа. Эксперт отмечает задание:

- ✓ неоднозначно,
- ✓ слабо однозначно,
- ✓ однозначно.

14. *Краткость* представления тестовой ситуации. Эксперту предлагается оценить задание как очень длинное, краткое или предельно краткое. Оказалось, что авторы БТЗ, чаще всего, по педагогическим и некоторым психологическим дисциплинам, первоначально использовали очень длинные формулировки, что сразу сказало на испытуемых и общих параметрах БТЗ. Тестирующиеся быстро уставали вчитываться в громоздкие формулировки задания, переставали задумываться над ответами, отвечали наобум.

15. *Свернутость* тестового задания предполагает такую формулировку, которую можно охватить одним взглядом. Предпочтение отдается рисункам, которые в себе содержат всю необходимую информацию и исключают специальную формулировку задачи. Этот показатель связан с ограничением времени, отведенным на организацию тестирования, когда испытуемый не имеет возможности долго вчитываться в текст.

16. *Ясность смысла* тестовой ситуации отличается от однозначности. Может оказаться ситуация, когда используется редкое понятие. Например, современные ученики школ и даже студенты вуза

не часто слышат «баба копра». Это однозначное понятие, используемое в задачах по физике, является неизвестным, следовательно, неясным для испытуемого.

17. *Доля* минимально подготовленных «*троечников*», которые смогут правильно выполнить тестовое задание (в процентах). Эта оценка ставится экспертом «на глазок», исходя из собственного опыта преподавания данной дисциплины. Она связана с коэффициентом успешности выполнения данного задания, поэтому этот показатель дублируется проведением пробных тестирований в испытательных лабораториях. Для правильной оценки необходимо представить некий средний класс, группу студентов; выделить из них только слабо подготовленных учеников – троечников и оценить в процентах, сколько из них сделают данное тестовое задание.

18. *Доля* минимально подготовленных «*отличников*», которые смогут правильно выполнить ТЗ (в процентах). Этот показатель оценивается аналогично предыдущему.

Два последних показателя позволяют в дальнейшем определить, так называемую, критериальную валидность теста. В тестологии различают содержательную и критериальную валидность. *Критериальная валидность* – характеристика теста, отражающая степень уверенности в том, что первичный тестовый балл реально определяет достижение определенного уровня знаний, умений или навыков. Критериальная валидность определяется на основании сравнения результатов тестирования и показателей, полученных независимым путем (например, путем традиционных оценок, экспертных суждений, результатов других тестов, валидность которых уже установлена). Существуют достаточно простые методики расчета критериального значения теста на основе двух последних показателей экспертизы тестовых заданий, например, метод Аббея.

При анализе более 100 протоколов экспертных заключений, сделанных учителями физики по совокупности тестовых заданий, предлагаемых для диагностики предметной обученности, последние показатели (17-й и 18-й) выявили яркую проблему, особенно после сопоставления с результатами выполнения данных диагностических работ их учениками. Выявилось огромное несоответствие мнения учителей об успешности выполнения и

дания «троечниками» и «отличниками» и реальным выполнением работы учениками. Оказалось, что учителя *всегда* завышают этот параметр. Педагоги считают, что больше «троечников» и, практически, все «отличники» должны сделать тестовые задания.

Проведение столь объемной экспертизы, включение более 100 экспертов продиктовали необходимость определения критериального значения (критериальной валидности). Проблема задания критериального значения относится к тесту и не является обязательной для БТЗ. Однако именно экспертная оценка некоторых параметров каждого тестового задания позволяет получить характеристику каждого теста, составленного на основе БТЗ. Более того, возможна обратная процедура: по данным показателям (двум последним в перечислении оцениваемых параметров) можно подобрать равноценные диагностические варианты или скомпоновать целую совокупность тестов с заранее определенными характеристиками. К сожалению, пока возлагаемые на данные параметры надежды не оправдались. Это можно объяснить некомпетентностью экспертов – они являются специалистами в своей области знаний, но не в области тестологии. Этот пример еще раз подчеркивает важность подбора и подготовки экспертов.

Опыт анализа экспертных заключений показал, что весьма желателен еще один показатель – правильный ответ на задание. Нельзя воспринимать этот показатель как проверку знаний эксперта, а следует считать его еще одним уровнем «чистки» тестовых заданий от ошибок и неточностей.

По неоднозначным заданиям эксперт может высказать особое мнение в примечании (рецензии) к таблице, в которой желательно стандартизированными значками оценить тестовое задание по всем параметрам.

Конечно, при проведении внутренней предварительной экспертизы может быть уменьшено количество параметров. Однако при проведении внешней экспертизы, особенно при сертификации БТЗ по всем перечисленным параметрам, эксперт высказывает свое мнение – это является одним из требований к оформлению документов на сертификацию.

Экспертиза тестового задания заканчивается фиксацией оценок каждого тестового задания по всем перечисленным параметрам.

Экспертиза банка тестовых заданий (БТЗ) включает в себя как элемент экспертизу каждого тестового задания, однако имеет и дополнительные параметры, оценивающие в целом весь БТЗ. В качестве самостоятельных показателей оценки БТЗ выступают:

- соответствие всего банка государственному образовательному стандарту или рабочей программе по дисциплине, отсутствующей в стандарте, например, элективные курсы, спецкурсы по выбору и т.п.;

- методическое обеспечение, полнота документов и структурных элементов;

- достаточность тестовых заданий на каждый уровень БТЗ. При возможности в одном БТЗ скомпоновать множество тестов для исключения узнавания их, особенно для проведения адаптивного тестирования, ведь БТЗ должен быть избыточен. А именно на каждый конечный уровень деления «дерева БТЗ» (тема, или понятие, или дидактическая единица) должно быть такое количество заданий, которого достаточно для случайного выбора и адаптации к уровню тестируемого. Однозначно такое количество задать трудно;

- полнота охвата всех дидактических единиц государственного образовательного стандарта или содержательная валидность БТЗ. Под валидностью теста понимают характеристику, показывающую, в какой мере тест измеряет именно то качество, для измерения которого он создан. Чем выше валидность теста, тем более обосновано использование результатов тестирования для выводов и предсказаний. Данная характеристика, прежде всего, относится к тестам, но может применяться и к банкам тестовых заданий (БТЗ), на основе которых может быть составлена целая совокупность тестов.

Содержательная валидность теста – характеристика теста, определяющая степень уверенности, что задания достаточно полно отражают содержание определенной области знания, владение всеми навыками, наличие которых измеряет тест, при этом тест не проверяет наличия второстепенных или ненужных знаний, умений и навыков. Другими словами, содержательная валидность – это характеристика, определяющая соответствие тестовых зада-

ний содержанию образовательного стандарта или учебной программе. Содержательная валидность оценивается с помощью метода экспертных оценок.

О содержательной валидности БТЗ судят по спецификации и методам, которые использовались при разработке банка. Для составления заключения эксперт анализирует:

- ✓ какова процедура отбора заданий, обеспечивает ли она включение нужного материала и возможность судить по выполнению заданий об уровне знаний испытуемого в данной области и овладении им специфическими и для данной области навыками;
- ✓ достаточно ли ясно указано владение какими именно знаниями или навыками выявляет БТЗ;
- ✓ все ли основные элементы содержания данной дисциплины представлены заданиями в БТЗ;
- ✓ достаточно ли равномерно для каждой дидактической единицы распределены задания по формам ТЗ, количеству, трудности;
- ✓ не включены ли задания, относящиеся к темам, которые не изучались.

Эксперт выставляет оценку содержательной валидности теста в диапазоне от 0 до 1. Единица выставляется, если БТЗ полностью валиден по содержательности. БТЗ считается содержательно-валидным, если значение оценки содержательной валидности составляет больше 0,8. Если эксперт оценивает содержательную валидность БТЗ меньше 1, то должен высказать замечания по содержательной валидности в «Лист замечаний» экспертизы.

Содержательную валидность можно оценить даже количественно, если полноценно проведена экспертиза каждого тестового задания, а именно определен параметр соответствия стандарту. По нашему мнению, она может быть грубо определена как отношение количества тестовых заданий, оцененных экспертом как полностью соответствующих государственному образовательному стандарту, к общему количеству заданий.

Положительное заключение в целом по БТЗ можно сделать, если все дидактические единицы образовательного стандарта представлены в структуре БТЗ, превышение над стандартом составляет не более 15%.

Для выделения качеств, которые могут быть оценены данным БТЗ, необходимо определить область использования и сферу применимости банка. Для этого БТЗ должен сопровождаться пояснениями о рекомендуемой цели применения. Эти пояснения могут включать описание совокупности тестируемых (испытываемых), на которых ориентирован БТЗ, например, возраст тестируемых, определенный год обучения, уровень способностей и т.п.

Эксперт должен оценить, соответствует ли БТЗ поставленным целям, ориентирован ли он на выбранную разработчиками группу обучающихся, и отразить это в протоколе экспертизы.

Правильно организованный банк тестовых заданий в своем составе содержит семейство программно-дидактических *тестов* (ПДТ). Именно по этим тестам в исследовательской лаборатории студенты проходят пробное тестирование. По результатам тестирования делается заключение о точности измерения данным БТЗ.

Тесты, как и банк тестовых заданий, являются совокупностью тестовых заданий, поэтому все параметры, по которым проходит экспертиза БТЗ в целом, справедливы для отдельных тестов. Однако для разных тестов, сформированных на основе одного и того же банка тестовых заданий, некоторые характеристики могут различаться. Очевидно, что неправильно сконструированный тест может иметь очень низкую содержательную валидность при высокой содержательной валидности всего БТЗ. Это возможно, если тест составляется с большим пренебрежением к проверяемым дидактическим единицам. Как говорилось выше, критериальная валидность — это специфическая характеристика теста, и определяется она именно для теста, а не всего БТЗ. Реальное тестирование проводится по тестам, поэтому все количественные статистические показатели, за исключением коэффициента успешности выполнения конкретного задания, рассчитываются для теста.

Кроме этого, проведение тестирования позволяет существенно дополнить качественные показатели количественными, вычисленными с использованием анализа данных методами математической статистики. Однако проблемы определения и интерпретации количественных характеристик теста являются темой отдельного разговора.

PROBLEM OF QUALITY OF TESTING MATERIALS

A.N. Velichko, O.A. Efremova

In the paper approaches to study the quality of testing materials are disclosed, significance of diagnostic data examination is determined, the problems displayed at practical examination of a test task bank are presented.

Key words: diagnostics, examination, tasks of testing character, tests.

ВЕС ТЕСТОВОГО ЗАДАНИЯ И МЕРА СТРУКТУРИРОВАННОСТИ ЗНАНИЙ – ЧТО ОБЩЕГО?

В.В. Кромер

При проведении педагогических измерений модель измерения должна соответствовать объекту измерения, в противном случае – существенно снижается валидность тестовых результатов. Вес задания и мера структурированности знаний испытуемого – параллельные характеристики при отнесении их к соответствующей стороне тестирования.

Ключевые слова: уровень подготовленности испытуемого, мера трудности задания, дифференцирующая способность задания, мера структурированности знаний, модель Раша, модель Бирнбаума.

В настоящее время тестирование осуществляется в рамках классической модели тестирования, развитой в первой половине XX в., и современных моделей тестирования, называемых в англоязычной литературе IRT-моделями. В.С. Аванесов [1, с. 24 – 25] пишет: «Имеется несколько вариантов одинаково неудовлетворительных переводов названия IRT на русский язык», и использует в цитируемой работе термин «математико-статистическая теория оценки латентных параметров заданий и уровня подготовленности испытуемых». Специалистами Федерального Центра тестирования РФ используется термин «Теория моделирования и параметризации педагогических тестов (ТМПТТ)» [9, с. 7].