

КАРАВАЕВА Ирина Александровна

ВАЛИДНОСТЬ ПЕДАГОГИЧЕСКИХ ТЕСТОВЫХ ИЗМЕРИТЕЛЕЙ
КАК СРЕДСТВО ПОВЫШЕНИЯ ОБЪЕКТИВНОСТИ
КОНТРОЛЯ ОБУЧЕННОСТИ

13.00.01 - общая педагогика, история педагогики и образования

Автореферат

диссертации на соискание ученой степени
кандидата педагогических наук



Работа выполнена в Институте повышения квалификации и переподготовки работников образования Удмуртской Республики

Научный руководитель: доктор педагогических наук,
профессор Черепанов В.С.

Официальные оппоненты: доктор педагогических наук,
профессор Мирошниченко А.А.

кандидат педагогических наук,
доцент Снигирева Т.А.

Ведущая организация Вятский государственный
гуманитарный университет

Защита диссертации состоится «14» ноября 2003г. в 11-30 час.
на заседании диссертационного совета Д 212.275.02 по защите диссертаций на соискание ученой степени доктора педагогических наук по специальности] 3.00.01 - общая педагогика, история педагогики и образования в Удмуртском государственном университете по адресу." 426034, г. Ижевск, ул. Университетская 1, корпус 6.

С диссертацией можно ознакомиться в библиотеке Удмуртского государственного университета (г. Ижевск, ул. Университетская 1, корпус 2).

Автореферат разослан «9» октября 2003 г.

Ученый секретарь
диссертационного совета,
кандидат психологических наук



Э.Р. Хакимов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность исследования. К важнейшим задачам модернизации образования относятся обеспечение системного реформирования содержания образования, дальнейшая разработка и внедрение государственных стандартов для всех ступеней и форм образования, развитие Национальной системы оценки качества образования, и на этой основе создание условий для объективного контроля качества знаний обучающихся.

Проблемам качества образования, методологии педагогического контроля знаний учащихся общеобразовательных школ и студентов посвящены труды В.С. Аванесова, С.И. Архангельского, В.П. Беспалько, З.Д. Жуковской, И.Я. Лернера, М.Н. Скаткина, Н.Ф. Талызиной и других исследователей. Проблемам формирования системы объективной оценки качества обучения посвящены материалы ряда Всероссийских симпозиумов «Квалиметрия человека и образования: методология и практика» (г. Москва, 1994-2002).

Возникшая на сегодняшний день необходимость обеспечения единого образовательного пространства России потребовала введения единого государственного экзамена в тестовой форме в масштабе всей страны. В этой связи особое внимание уделяется проблемам разработки и использования качественных и объективных тестовых измерителей. В традиционных методиках проверки качества обученности учащихся, как правило, отсутствуют четкие объективные критерии оценки, преобладает субъективный характер оценочных суждений.

Анализ научной литературы позволяет утверждать, что объективная оценка качества обучения возможна при использовании, прежде всего, валидных и надежных педагогических тестовых измерителей. Однако, процедуры валидации педагогических тестов, встречающиеся в научной литературе, строятся по аналогии с более сложными инструментами психолого-педагогических и социологических исследований, объекты которых латентны и слабо формализуемы. Объективность результатов таких исследований подтверждается через привлечение сложного математико-статистического аппарата, использование которого в рамках текущего дидактического контроля нетехнологично. Для рубежного и итогового контроля тестологи предлагают использовать сертифицированные тесты, построенные на основе методов теории моделирования и параметризации педагогических тестов, которая является основой для построения надежных тестовых измерителей результатов обучения. Валидность таких тестов базируется на статистических характеристиках тестовых заданий, зависящих от свойств выборки стандартизации, что не вполне обеспечивает выполнение требований образовательных стан-

дартов по учебному предмету и не гарантирует сохранение единого образовательного пространства.

Степень разработанности проблемы. Общая проблема валидности эмпирического психолого-педагогического исследования сложных латентных объектов отражена в трудах А. Анастаси, К.М. Гуревича, К. Ингенкампа, В.А. Дюка, Т. Кука, Д. Кэмпбелла, Я. тер Лаака, Р. Линна и других авторов.

Различные аспекты создания качественных педагогических тестовых измерителей и проведения с их помощью объективной оценки результатов обучения нашли свое отражение в работах А.Н. Майорова, Ю.М. Неймана, В.А. Хлебникова, М.Б. Челышковой и других исследователей, а также широко обсуждались на Всероссийских конференциях «Развитие системы тестирования в России» (г. Москва, 1998-2002);

Существенный вклад в решение проблемы валидности педагогических тестов внесли В.С. Аванесов, А.С. Казаринов, А.Н. Майоров, Е.А. Михайлычев. Ее решение в той или иной мере разработано в диссертационных исследованиях Е.Н. Лебедевой, В.А. Шухардиной и других. Однако многие исследователи не всегда используют теоретически обоснованные процедуры определения валидности педагогических контрольных материалов (в том числе и тестов), не определяют полную систему валидности для них. В связи с этим, возникает **противоречие** между растущей потребностью использования в учебном процессе качественных тестов, объективно контролирующих результаты обучения, и недостаточной разработанностью процедуры определения валидности для обеспечения их качества.

Указанное противоречие порождает проблему исследования, заключающуюся в обосновании необходимой и достаточной системы видов валидности и разработке квалиметрических процедур определения валидности педагогических тестов и тестовых заданий для обеспечения объективности педагогического контроля обученности.

Цель настоящего исследования состоит в разработке и обосновании системы валидности педагогических тестовых измерителей для обеспечения объективности контроля обученности учащихся.

Объект исследования - процесс педагогического контроля обученности учащихся на основе тестовых технологий.

Предмет исследования - система валидности педагогических тестовых измерителей.

Общая гипотеза исследования основана на предположении о том, что достижение объективности тестового контроля результатов обучения обеспечивается качеством инструментария педагогического контроля, опирающегося на обоснованную систему валидности теста.

Производной от общей гипотезы являются следующие частные предположения о том, что:

1) внутренним фактором обеспечения объективности тестовых результатов измерения обученности является ограниченная и полная система валидности педагогического теста:

- полнота системы валидности достигается обеспечением внутренней валидности теста;
- условие ограниченности системы валидности зависит от характера объекта измерения;

2) внешним фактором обеспечения объективности результатов педагогического контроля является технологичная процедура валидизации педагогического теста, основанная на методе экспертных оценок.

В соответствии с предметом, целью и гипотезой исследования определены следующие задачи:

1. Выявить характеристики видов валидности педагогического теста в аспекте типологизации объектов педагогического контроля и объективности тестовых результатов.

2. Обосновать способы определения системы валидности педагогических тестов.

3. Определить критерии для ваяидизации тестовых измерителей обученности.

4. Установить показатели качества теста для обеспечения объективности результатов при педагогическом контроле обученности.

5. Провести экспериментальное исследование разработанной процедуры определения валидности тестовых измерителей.

Теоретико-методологическая основа исследования определяется поставленными целями и задачами и строится на разработанных теориях и методах педагогических исследований в следующих областях:

- содержания образования и стандартов в образовании (В.И. Байденко, М.В. Кларин, В.В. Краевский, И.Я. Лернер, В.С. Леднев, М.Н. Скаткин, В.С. Цетлин),
- качества результатов эмпирического исследования (А. Анастаси, Д. Кэмпбелл, А.С. Казаринов, Я. тер Лаак, К. Левин, А.А. Мирошниченко, Е.А. Михайлычев, А.И. Самыловский),
- педагогического контроля и квалиметрии (С.И. Архангельский, В.П. Беспалько, П.Я. Гальперин, Б.С. Гершунский, З.Д. Жуковская, В.А. Кальней, В.И. Михеев, В.П. Панаскж, М.М. Поташник, Н.А. Селезнева, А.И. Субетто, Н.Ф. Талызина, В.С. Черепанов, С.Е. Шишов),
- образовательной тестологии (В.С. Аванесов, А.Н. Майоров, Ю.М. Нейман, О.А. Татур, В.А. Хлебников, М.Б. Чельшкова).

Ввиду того, что в педагогической тестологии пока нет однознач-

ных трактовок ряда определений и терминов, в работе используется следующий терминологический аппарат: *система валидности педагогического теста (общая валидность теста)*, понимаемая как комплексная характеристика, состоящая из четырех основных видов (теоретической, содержательной, эмпирической и валидности теста по форме) и двух типов (внутренней и внешней).

К внешнему типу валидности относится вид валидности, направленный на обеспечение экстраполяции тестовых результатов, полученных на выборке нормирования, на всю генеральную совокупность тестирующихся. *Внутренняя валидность* предназначена для обеспечения связи независимых и зависимых переменных, причины и эффекта эксперимента. *Валидизация* трактуется как качественная процедура определения валидности теста, основанная на методе экспертных оценок.

Теоретическая валидность представлена *валидностью конструкта*, или *конструктивной валидностью*, которая задает адекватное объекту теоретическое представление предмета измерения. *Содержательная валидность* отражает адекватность соответствия содержания теста теоретическому конструкту. Подвидом содержательной валидности является внешняя *куррикулярная валидность*, отражающая степень соответствия содержания теста структурному теоретическому конструкту. *Эмпирическая валидность* направлена на эмпирическое подтверждение гипотезы исследования. К ее подвидам относятся: *прогностическая, критериальная, сопоставительная, соответствия, текущая валидность*, а также *факторная* и ряд других. К *валидности теста по форме* относится *техническая*, или *композиционная валидность*, предназначенная для установления соответствия структуры и формы теста принятым в тестологии нормам.

Тип валидности зависит от характера объекта измерения. Под *латентным объектом педагогического контроля* в педагогическом тестировании предлагается понимать уровень обученности. *Структурный теоретический конструкт* определяется как репрезентативное поэлементное представление контролируемого содержания учебного предмета. Под *открытым объектом педагогического контроля* понимается результат обучения (обученность). Обученность представляется *целостным теоретическим конструктом* предметного содержания, который определен структурой предмета и ограничен теоретическими подходами к его изложению.

Методы исследования. Для решения поставленных в исследовании задач и проверки гипотезы была использована совокупность общенаучных и педагогических методов: *теоретические* (изучение и анализ философской, психолого-педагогической и методологической литературы по проблеме; системный подход, аналогия, моделирование, гипоте-

тико-дедуктивный метод конструирования объекта измерения); *эмпирические* (тестирование, метод групповых экспертных оценок); *математико-статистические* (обработка эмпирических данных).

Этапы исследования. На первом этапе - теоретико-поисковом (1998-1999 гг.) — осуществлялся анализ отечественной и зарубежной научной литературы, диссертационных работ по теме исследования. В результате были сформулированы проблема и гипотеза исследования, разработана программа исследования.

На втором этапе — опытно-экспериментальном (2000-2001 гг.) - конструировались и апробировались тестовые измерители обученности, проводилась разработка и корректировка процедуры определения валидности тестовых измерителей обученности.

На третьем этапе - заключительном (2002-2003 гг.) - проводился анализ и обобщение полученных результатов экспериментального исследования, определялась степень их объективности и осуществлялось внедрение валидных тестовых измерителей в учебный процесс ряда общеобразовательных учреждений Удмуртской Республики.

Экспериментальная база исследования — подготовительное отделение Ижевского государственного технического университета.

Научная новизна исследования заключается в следующем:

- выявлено, что для *латентного* объекта исследования основной характеристикой качества теста является его надежность, а внешняя валидность носит вероятностный характер и не может ограничиваться конечным числом ее видов;

- установлено, что для *открытого* объекта исследования (обученности) содержательная валидность теста является внутренней, что и позволяет обеспечить объективность тестовых результатов;

- определены специфические способы технологизации системы валидности педагогических тестов, к которым относятся: 1) выбор критерия валидизации, опирающийся на содержание теста, и 2) установление внутренней валидности теста на основе межличностного соглашения экспертов.

Теоретическая значимость исследования состоит в том, что:

- представлена зависимость объективности тестовых результатов в аспекте их соответствия определенному сконструированному объекту измерения от внутренней валидности тестового измерителя обученности;

- выявлены и обоснованы условия технологизации внутренней валидности целостного конструкта предметного содержания;

- обоснована относительно полная система валидности педагогического теста для открытого объекта педагогического контроля.

Практическая значимость исследования заключается:

- в разработке квалиметрической процедуры конструирования валидных тестовых измерителей обученности;
- в создании на основе данной технологии тестовых материалов, которые применяются для дидактического контроля обученности на подготовительном отделении Ижевского государственного технического университета и в ряде общеобразовательных учреждений Удмуртской Республики.

Предложенная процедура определения валидности педагогических тестов может быть использована при оценке текущих и итоговых знаний учащихся по математике и другим предметам в образовательных учреждениях различного типа.

На защиту выносятся следующие положения:

1. Для *латентного* объекта педагогического контроля (уровня обученности) основной характеристикой качества теста, построенного на основе структурного теоретического конструкта предметного содержания, является надежность; внешняя валидность носит вероятностный характер и не может ограничиваться конечным числом ее видов, а внутренняя валидность отсутствует.

2. Для *открытого* объекта педагогического контроля (обученности) содержательная валидность теста, построенного на основе целостного теоретического конструкта предметного содержания, является внутренней валидностью, что позволяет установить соответствие тестовых результатов данному сконструированному объекту контроля.

3. Для *латентного* объекта педагогического контроля система валидности педагогического теста состоит из внешних видов валидности: конструктивной, куррикулярной, критериальной. Для *открытого* объекта педагогического контроля система валидности педагогического теста определяется внутренними видами валидности: конструктивной, содержательной, критериальной. Обе системы дополняются независимой относительно ее типа технической валидностью.

4. Содержание теста как целостная модель контролируемого предметного содержания является внутренним критерием его валидации и опирается на принцип межличностного соглашения экспертов.

5. Объективность тестовых результатов достигается за счет обеспечения внутренней содержательной валидности теста и исключения субъективных факторов из процесса построения теста.

Апробация результатов исследования осуществлялась на научных конференциях и семинарах, а также в виде публикаций.

Основные положения и результаты исследования обсуждены на девятом и десятом Всероссийских симпозиумах по квалиметрии человека и образования (г. Москва, 2000-2002 гг.); Всероссийских конференци-

ях «Развитие системы тестирования в России» (г. Москва, 1999 г., 2002 г.), 5-й Российской университетско-академической научно-практической конференции (г. Ижевск, 2001 г.), региональной научно-практической конференции «Развитие системы тестирования в Удмуртской Республике» (г. Ижевск, 2001 г.), и других научно-методических конференциях. Исследование входило в план научно-исследовательских работ Уральского отделения Российской академии образования по комплексной программе «Образование в Уральском регионе: научные основы развития и инноваций» на 2001-2005 гг. (раздел 2 тема № 29).

Достоверность результатов исследования и обоснованность научных положений и выводов обеспечиваются последовательной реализацией методологической основы процесса исследования на всех его этапах; соответствием теоретических и эмпирических методов исследования предмету исследования и поставленным задачам; соответствием полученных данных положениям педагогической теории и практики в области образования.

Структура и объем диссертации.

Работа состоит из введения, трех глав, заключения, списка литературы, приложений. Общий объем текста — 173 страницы, работа включает 10 рисунков, 7 таблиц, 5 приложений.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы исследования, рассматриваются основные характеристики научного аппарата: цель, объект и предмет исследования, его гипотеза и задачи; характеризуются методы и определена теоретико-методологическая основа исследования, новизна результатов, их теоретическая и практическая значимость, сформулированы защищаемые положения.

Глава первая «Система валидности педагогического теста» посвящена анализу современного состояния проблемы обеспечения качества тестирования как экспериментального педагогического исследования; объективности тестовых выводов, их соответствия объекту исследования.

Проблема построения системы валидности эмпирического исследования и обеспечения качества его результатов обсуждались в работах А. Анастаси, В.А. Дюка, Т. Кука, Д. Кэмпбелла, Я. тер Лаака, К. Левина, Р. Линна, Д. Мангейма, Р. Рич, Э. Мэнстида, Г. Симины.

В научной литературе по педагогическому тестированию (В.С. Аванесов, А.Н. Майоров, Ю.М. Нейман, В.С. Хлебников, М.Б. Чельшкова) объективность тестовых результатов сводится к контролю над ус-

ловиями проведения эксперимента на различных его этапах. Это — объективность: 1) проведения эксперимента, 2) обработки данных, 3) интерпретации (К. Ингенкамп). Обозначенные условия определяются как необходимые для получения достоверных или валидных выводов и являются внешними по отношению к объекту эмпирического исследования. Объективность как соответствие результатов исследования его объекту авторами не рассматривается. Это связано прежде всего со сложностью, неопределенностью, то есть латентностью объекта исследования.

В педагогике, как и в других социальных науках, как правило, объекты исследования являются латентными, то есть недоступными для непосредственного измерения. Общность характеров объектов позволяет педагогам заимствовать методы исследования из смежных наук, таких, как социология и психодиагностика, в которых были разработаны статистические методы подтверждения достоверности гипотез.

В диссертационной работе вводится достаточное условие объективности - соответствие тестовых результатов объекту измерения через обеспечение внутренней содержательной валидности инструмента исследования (теста) на основе открытого и определенного, а не латентного объекта исследования; доказываемая технологичность процедуры определения валидности теста, измеряющего открытый объект - обученность; строится ограниченная и полная система валидности педагогического теста, контролирующего обученность.

В педагогической тестологии качество тестовых результатов непосредственно связывают с качеством инструментария контроля результатов обучения (дидактического теста). Анализ валидности педагогического теста, построению ее системы и обоснованию различных ее видов посвящены работы по проблемам педагогических измерений В.С. Аванесова, К. Ингенкампа, А.С. Казаринова, А.Н. Майорова, Е.А. Михайлычева, А.И. Самыловского и других исследователей.

В теоретической части работы раскрывается зависимость системы видов валидности теста от выбранного принципа истинности (Я. тер Лаак) и типа объекта исследования: латентного или открытого. Как правило, система состоит из трех основных видов: **1) валидности конструкта (теоретической), 2) содержательной, 3) прогностической (эмпирической)** валидности. Показано, что иные названия валидности, встречающиеся в литературе, являются подвидами выше перечисленных. Тест обладает валидностью конструкта, если теоретический конструкт адекватно представляет объект исследования.

В теоретическом конструкте обычно выделяются независимые переменные. Управление ими в ходе эксперимента приводит к экспериментальному эффекту, который выражается количественно через фикса-

цию в результатах эксперимента посредством измерения зависимых переменных. Через **содержательную** валидность контролируется репрезентативность отображения содержанием теста теоретического конструкта. Содержание теста в этом случае связывает независимые и зависимые переменные, причину и эффект. Эти виды валидности выражаются качественными характеристиками (валиден - невалиден), как правило, с применением метода экспертных оценок. Экспертами оценивается, в какой степени теоретический конструкт репрезентативно представляется тестовым содержанием. Количественное подтверждение достоверности результатов представляет **эмпирическая** валидность, которая выражается через корреляционный коэффициент

Анализ полноты соответствия тестового содержания теоретическому конструкту отсутствует. Неполнота представления тестовым содержанием теоретического конструкта ставит под сомнение внутреннюю связь причины и эффекта, которая требует эмпирического подтверждения. В работе теоретически показано, что для латентного объекта невозможно однозначное установление связи между ним и тестовыми результатами через содержательную валидность теста ввиду сложности объекта и ограниченной длины теста. С этой целью вводится понятие **внутренней и внешней** валидностей.

Для установления **внутренней** валидностей в ходе эксперимента необходимо контролировать всевозможные внешние факторы, ей угрожающие (Д. Кэмпбелл), количество которых неопределенно. Это приводит к системе валидности теста, содержащей бесконечное количество разных видов (А.С. Казаринов). Система валидности становится неограниченной, полная валидизация невозможна, что делает процесс определения валидности теста нетехнологичным

Внешняя валидность позволяет экстраполировать тестовые результаты на группу тестирующихся большую, чем экспериментальная (Д. Кэмпбелл) или на содержание большее, чем тестовое. Наличие внешней валидности инструментария не является достаточным условием объективности тестовых результатов.

На основе анализа определений педагогического теста для контроля результатов обучения (К. Ингенкамп, В.С. Аванесов, А.Н. Майорова, Е.А. Михайлычева, Ю.М. Нейман, В.А. Хлебников) в работе обоснована его система валидности, состоящая из четырех видов валидности: конструктивной, содержательной, критериальной и технологической, под которой понимается соответствие формы теста принятой в дидактической тестологии норме. Такая система валидности теста достаточна тогда и только тогда, когда содержательная валидность является внутренней.

Результат проведенного исследования показал, что при измерении

латентного объекта, под которым в дидактической тестологии понимается уровень обученности (Ю.М. Нейман, В.А. Хлебников, М.Б. Чельшкова и др.), эти виды валидности педагогического теста относятся к внешнему типу валидности. Отсутствие внутреннего типа валидности такого теста не позволяет обеспечить контроль над тем, что именно измеряет тест. На первое место в иерархии значимости показателей качества теста выходит надежность инструмента.

В исследовании показана зависимость внешнего критерия эмпирической валидации от цели тестирования и выбранного принципа истинности (принцип когеренции, функциональности или консенсуса). Этот критерий обуславливает различия в интерпретации тестовых результатов и относит тест к одному из двух видов: нормативно- или критериально-ориентированному. Педагогический тест относится к нормативно-ориентированным при условии, что в качестве такого критерия эмпирической валидации выбирается внутренняя норма, а именно, нормальный закон распределения величин для выборки стандартизации. Валидность такого теста обоснована принципом функциональности, что позволяет ранжировать тестирующихся по степени владения тестовым содержанием. Главной характеристикой качества становится точность как степень свободы тестовых результатов от ошибки измерения, или надежность теста.

Выбор в качестве внешнего критерия валидации теста иных результатов измерения того же объекта, валидность которого всегда окзывается под вопросом, классифицирует его как критериально-ориентированный. В работе показано, что прогностическая или критериальная валидность результатов тестирования в процессе эмпирической валидации понимается как их соответствие результатам другой внешней и не более валидной оценки; содержательная валидность носит характер внешней валидности ввиду его ограниченности, а основным показателем его качества становится надежность в смысле однородности отражения тестом измеряемого предметного содержания. Отсутствие внутренней валидности такого теста делает измерение относительным в аспекте неполного соответствия содержания теста измеряемой предметной области, а значит - нетехнологичным.

В результате теоретического анализа работ по теме исследования сделан вывод о том, что объективность результатов в аспекте их соответствия объекту измерения достигается через замену латентного объекта (уровня обученности) на открытый и определенный - обученность. Обученность мы определяем как результат обучения, заключающий в себе понимание системы предметного содержания как целого, представленного в разных теоретических аспектах или в одном из них. При контроле обученности предметом измерения служит содержание, определенное

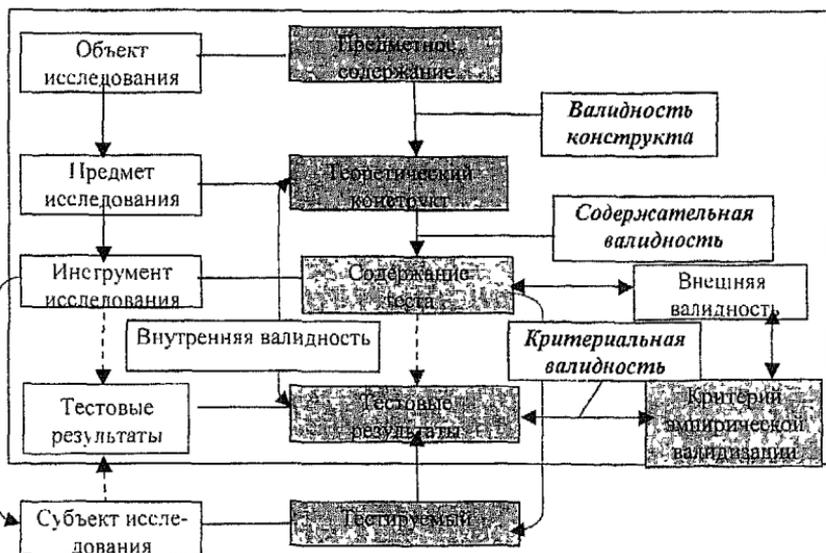


Рис. 1. Схема валидизации целостного конструкта *обученное*TM

структурой учебного предмета, ограниченное теоретическими подходами к его изложению и представленное для измерения в виде целостного теоретического конструкта. Открытый объект позволяет элиминировать субъект тестирования за пределы процесса валидизации, что обеспечивает ограниченность и полноту системы валидности теста, технологизацию процесса валидизации и получение объективных тестовых результатов в аспекте их соответствия сконструированному объекту измерения (рис. 1).

Глава вторая «Объективность контроля *обученное*TM как педагогическая проблема» посвящена поиску путей получения объективных тестовых результатов в аспекте их соответствия сконструированному объекту измерения через внутреннюю содержательную валидность теста; выбору критерия валидизации теста и принципа истинности, лежащего в основе валидности; построению на этой основе процедур определения валидности педагогического теста.

В работе показано, что технологичность теоретического конструкта зависит непосредственно от выбранного объекта измерения. Принятый в современной образовательной тестологии уровень *обученное*TM определяется структурным теоретическим конструктом, для которого характерно представление усвоенного знания в виде структуры пред-

метного содержания. Структурный теоретический конструкт отражен в спецификации теста как логическая структура предмета, подлежащая измерению. Анализ соответствующей литературы (Ю.М. Нейман, В.А. Хлебников, М.Б. Челышкова, А.Н. Майоров и др.) показал, что для полного измерения необходимо проверить знание всех элементов структуры предмета, что на практике не представляется возможным ввиду ограниченности теста. Валидность конструкта носит внешний характер. Содержание теста как репрезентативное поэлементное отражение структурного конструкта определяет содержательную валидность как куррикулярную. Зависимость качества тестовых заданий от субъективных свойств выборки стандартизации и неполнота представления содержанием теста теоретического конструкта делают куррикулярную валидность внешней. Система валидности теста со структурным теоретическим конструктом всегда оказывается неполной из-за отсутствия внутреннего типа валидности. Присутствие субъекта тестирования (выборки стандартизации) в процедуре определения валидности теста делает ее нетехнологичной.

Результат обучения - обученность, как открытый объект измерения, не требует проверки владения всеми элементами структуры предметного содержания. Достаточным для проверки **обученности** оказывается владение **теоретическими подходами**. Валидность конструкта оказывается внутренней. Содержание теста конструируется на основе заданий, моделирующих основные подходы к изложению содержания учебного предмета (числовой и функциональный подход для школьного курса математики), и является моделью теоретического конструкта. В работе на примере школьного курса алгебры и начал математического анализа проиллюстрировано построение теоретического конструкта посредством выделения двух теоретических подходов к изложению содержания курса: числового и функционального.

В данном исследовании предлагается построение содержания теста как модели теоретического конструкта, в которой измеряемое содержание представляется не как поэлементное отображение структуры, а как целостная модель предметного содержания. Смоделированное таким образом тестовое содержание адекватно представляет теоретический конструкт, а **содержательная** валидность теста становится **внутренней** и перестает быть куррикулярной. Для школьного курса математики целостная модель содержания теста строится посредством моделирующих тестовых заданий двух типов: числовых и функциональных.

Поскольку объект исследования перестает быть латентным, то исчезает необходимость привлекать внешний критерий для опосредованной эмпирической валидации. Критерий становится внутренним. Целостный теоретический конструкт определен структурой предмета, ог-

раничен теоретическими подходами предметного содержания. В то же время его модель (тестовое содержание) инвариантна относительно структуры предмета и логики его изложения, так как ее основаниями выступают наиболее общие, чем структура и логика предмета, характеристики - владение теоретическим подходом. Одновременно теоретический конструкт вариативен в аспекте наполнения модели конструкта предметным содержанием. Следовательно, существует неограниченное количество представлений теоретического конструкта в виде содержательной модели целостного предметного содержания. Каждая из таких моделей, в том числе и содержание самого теста, может выбираться в качестве внутреннего критерия валидизации, вследствие чего критериальная валидность становится внутренней, наряду с конструктивной и содержательной валидностью. Их внутренний характер обеспечивает объективность тестовых результатов в аспекте соответствия открытому объекту измерения.

Внешняя валидность тестов с целостным теоретическим конструктом рассматривается в аспекте экстраполяции результатов тестирования на содержание большее, чем тестовое, а надежность, направленная на более точное описание конструкта, представляется как однородность содержания теста: чем выше содержательная валидность теста, тем выше однородность его тестовых заданий, направленная на повышение точности представления теоретического конструкта.

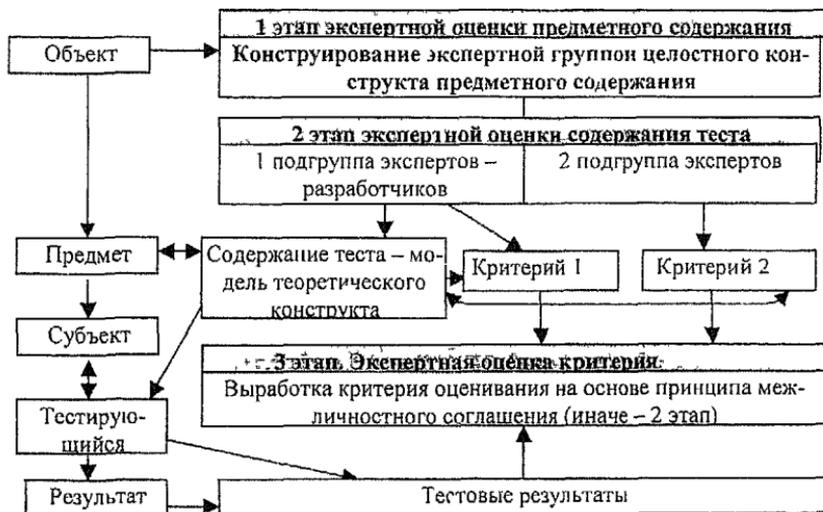


Рис.2. Блок-схема алгоритма валидизации педагогического теста с целостным конструктом

Техническая валидность - одинакова для обоих типов тестов. Она определяется экспертным методом посредством анализа тестовых заданий на соответствие их формы принятым в тестологии нормам.

Система валидности педагогического теста с целостным конструктом является полной. Она состоит из внутренних валидностей конструкта, содержания, критерия, а также технологической валидности.

Валидизация такого теста осуществляется экспертной группой на основе принципа консенсуса. Блок-схема процедуры создания валидного теста с целью объективного контроля результатов обучения осуществляется в соответствии с алгоритмом, представленным на рис. 2.

В третьей главе «Экспериментальная апробация процедуры определения валидности теста для измерения обученности» иллюстрируется конструирование валидного педагогического теста и проводится анализ объективности тестовых результатов проводимого на его основе эксперимента.

Технология валидизации педагогического теста, измеряющего обученность, включена в процедуру его конструирования и состоит из трех этапов (рис.2). Целью *первого* этапа является построение валидного теоретического конструкта контролируемого предметного содержания. На *втором* этапе конструируется содержание теста, обладающего внутренней содержательной валидностью и технической валидностью формы. Целью *третьего* этапа является установление валидного внутренне-

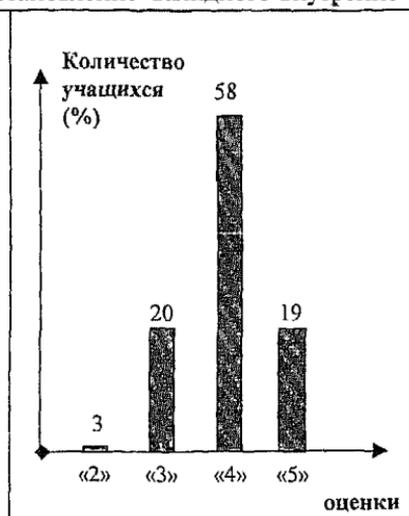


Рис.3. Диаграмма распределения частот выполнения заданий, контролирующей обученность

Рис. 4. Диаграмма распределения частот результатов централизованного тестирования

го критерия, основанного на содержании теста для измерения обученности.

В ходе экспериментального исследования, призванного подтвердить обоснованность теоретически созданной процедуры валидации тестовых измерителей, проводилось измерение обученности[™] учащихся подготовительного отделения Ижевского государственного технического университета (n=100 человек). Тестирование осуществлялось с помощью различных типов тестов. В качестве экспериментального теста с целостным теоретическим конструктом и открытым объектом измерения использовался тест, сконструированный на основе описанной процедуры. Результаты централизованного тестирования (ЦТ) за 2002 год выступали в роли контрольных тестовых показателей измерения уровня обученности - латентного объекта, На рис. 3 представлена диаграмма распределения частот результатов экспериментального тестирования в зависимости от суммы баллов, набранных за решение моделирующих заданий (от 0 до 5 баллов). По итогам экспериментального тестирования 9 слушателей набрали в сумме 5 баллов за задания, моделирующие функциональный подход. Они и были признаны владеющими знанием предмета как целым. На рис.4 изображена диаграмма распределения частот результатов контрольного централизованного тестирования той

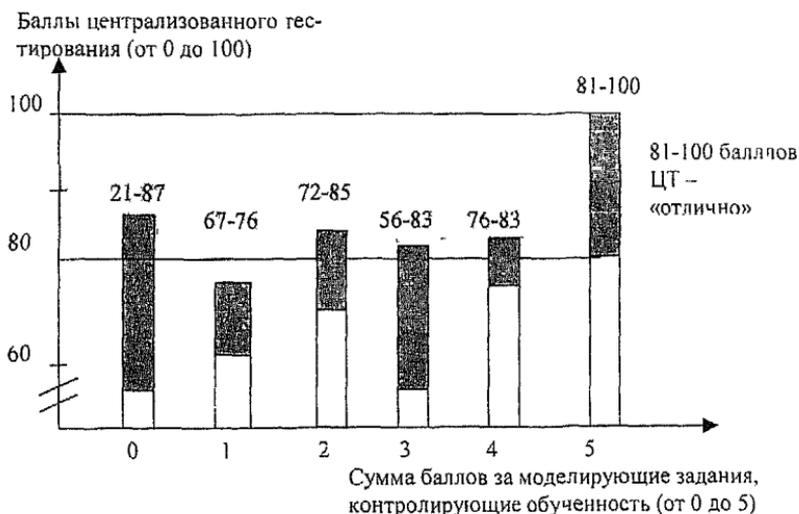


Рис.5. Диаграмма распределения баллов за централизованное тестирование в зависимости от суммы баллов за задания, контролирующие обученность

же группы слушате лей в зависимости от полученных отметок. На рис. 5 приведена диаграмма распределения баллов централизованного тестирования в зависимости от суммы баллов за моделирующие задания. В результате только 9 % учащихся показали знание предметного содержания как целого (рис.3), однако они вошли в группу из 19% учащихся, получивших отметку «отлично» в ходе ЦТ (рис.4). В ходе анализа результатов тестирований нами были сделаны следующие выводы: 1) прогноз о знании предмета как целого, сделанный на основании результатов тестирования, подтвердился в 100% случаев; 2) тесты с целостным конструктом демонстрируют высокую степень различительной способности, дифференцируют учащихся по степени владения содержанием предмета и являются более объективными.

В заключении диссертации подводятся основные итоги и намечаются перспективы дальнейшего исследования.

Результаты проведенной работы позволяют сделать следующие **выводы:**

1. В диссертации установлена необходимость регламентации использования качественных дидактических тестов в соответствии с целями педагогического контроля и типологизацией объекта измерения. Качественные тесты, построенные на основе структурного теоретического конструкта для измерения латентного объекта (уровня обученности), позволяют надежно ранжировать обучающихся по степени владения предъявленным тестовым содержанием; тесты с целостным теоретическим конструктом для измерения открытого объекта (обученности), обладающие внутренней содержательной валидностью, необходимо использовать при рубежном и итоговом контроле над результатами обучения.

2. В результате проведенного исследования обоснована система валидности педагогического теста для двух типов объекта контроля. Доказана невозможность построения ограниченной и полной системы валидности педагогического теста для измерения латентного объекта контроля; представлена концептуальная модель полной и ограниченной системы валидности тестового измерителя при педагогическом контроле обученности.

3. В работе обоснована зависимость объективности тестовых результатов от внутренней валидности тестового измерителя в аспекте их соответствия объекту педагогического контроля. Объективность тестового контроля обученности создает реальные перспективы обеспечения образовательного стандарта по предмету и сохранения единого образовательного пространства.

Выполненное исследование подтвердило правомерность выдвину-

тых в диссертации предположений и выводов. Внедрение его результатов в практику работы образовательных учреждений будет способствовать повышению объективности педагогического контроля результатов обучения.

В ходе работы над диссертацией были выявлены новые проблемы, требующие решения: 1) повышение качества обучения и интенсификации его процесса без увеличения объема структуры учебного предмета; 2) представление предметного содержания как целостного теоретического конструкта 3) отказ от обучения навыкам в пользу системных теоретических знаний по предмету; 4) углубленные исследования возможности и способов формирования теоретических знаний обучающихся.

Основное содержание диссертации отражено в публикациях:

1. Навалихина И.А. О технологии измерения адапционной валидности теста // Развитие системы тестирования в России: Тез. докл. Всерос. конф. Ч. 2. М.: Прометей, 1999. - С. 100-101.

2. Навалихина И.А. Содержательная валидность теста как основное условие объективности тестового контроля. // Развитие тестовых технологий в России: Тез. докл. Всерос. конф. М.: Центр тестирования МОРФ, 2002. С. 241-242.

3. Навалихина И.А. Проблема соотношения видов валидности в структуре общей валидности теста. // Квалиметрия человека и образования: методология и практика: Материалы IX Симпозиума / Под ред. Н.А. Селезневой, А.И. Субетто. М.: Исследовательский центр проблем качества подготовки специалистов, 2000.- С. 19-21.

4. Навалихина И.А. Соотношение внешней и внутренней валидности теста в теории ЕРТ. // Квалиметрия в образовании: методология и практика: Материалы X Симпозиума. Книга 2. М.: Исследовательский центр проблем качества подготовки специалистов, 2002. - С.85-87.

5. Навалихина И.А. Концепция структуры валидности педагогического теста. // Проблемы квалиметрии образования и дидактической тестологии: Сборник трудов направления «Измерения в педагогике». Ижевск: Изд-во ИжГТУ, 2003. - С. 88-91.

6. Навалихина И.А. Об эффективности использования нормативно-ориентированных и критериально-ориентированных тестов в аспекте валидизации теста // Развитие системы тестирования в Удмуртской Республике: Материалы регион, науч.-практ. конф. / Под ред. В.С. Черепанова. Ижевск: Изд-во ИжГТУ, 2001. - С.68-69.

7. Навалихина И.А. Мониторинг качества обученности при использовании тестовой формы контроля обученности // 5-я Российская университетско-академическая науч.-практич. конф.: Тез. докл. Ч.4. / Под ред. В.А. Журавлева, С.С. Савинского. Ижевск: Изд-во УдГУ, 2001 , - С. 33-35.

8. Навалихина И.А. Проблема соотношения трудности и сложности тестовых заданий при измерении качества обучения в аспекте валидности // Cogito. Вып. 5. Психолого-педагогические исследования профессиональной деятельности: Сб. ст./ Под ред. С.Ф. Сироткина. - Ижевск: Изд-во УдГУ, 2001.-С.123-130.

9. Навалихина И.А. Концепция структуры валидности педагогического теста // Проблемы квалиметрии образования и дидактической тестологии: Сб. трудов направления «Измерения в педагогике». - Ижевск: Изд-во ИЖГТУ, 2003. - С. 88-91.

10. Караваева И.А. Сравнительный анализ результатов тестирования на основе структурного и целостного конструкторов // Актуальные проблемы образования в высшей школе: Сб. материалов науч.-метод. конф. УдГУ / Под ред. проф., докт. пед. наук Г.С. Трофимовой. - Ижевск: Изд-во УдГУ, 2003. - С. 88-93.

11. Навалихина И.А. Проблема формирования базисного образовательного стандарта по математике // Удмуртия накануне третьего тысячелетия: Тез. докл. науч.-практ. конф. - Ижевск: Изд-во УдГУ, 1998. - 4.2, С.125.

12. Навалихина И.А. Проблема соотношения объективности, валидности и надежности в тестовом контроле с применением метода ГЭО // Жизнь в аспекте экологии и биологии: Материалы межд. совещ. «Жизнь и факторы биогенеза». - Ижевск: Изд-во УдГУ, 1999. - С. 71-73.

13. Караваева И.А. Целостный конструктор предметного содержания как предмет дидактического контроля обученности // Актуальные проблемы образования в высшей школе: Сб. материалов науч.-метод. копф. УдГУ / Под ред. проф., докт. пед. наук Г.С. Трофимовой. - Ижевск: Изд-во УдГУ, 2003. - С. 95-98.

В авторской редакции

Подписано в печать 7.10.2003. Бумага офсетная. Формат 60x84/16.
Уч. изд. № 1/сл. печ. п. - 1, III Тираж ЛР? экз. Заказ № 324

Типография Издательства ИЖГТУ.
426069, г. Ижевск, Студенческая, 7