

# Руководство по разработке тестовых заданий и конструированию педагогических тестов<sup>1</sup>

Автор-составитель Краснов Ю.Э.<sup>2</sup>

<b>ВВЕДЕНИЕ. ЭТАПЫ РАЗРАБОТКИ ПЕДАГОГИЧЕСКОГО ТЕСТА: ОТ ЗАДАНИЙ В ТЕСТОВОЙ ФОРМЕ ЧЕРЕЗ ТЕСТОВЫЕ ЗАДАНИЯ К ТЕСТУ</b> .....	2
<b>РАЗДЕЛ I. ЧТО НАДО СДЕЛАТЬ ДО РАЗРАБОТКИ ЗАДАНИЙ В ТЕСТОВОЙ ФОРМЕ И ПОСТРОЕНИЯ ТЕСТА? ДИАГНОСТИЧНОЕ ОПИСАНИЕ ЦЕЛЕЙ ОБУЧЕНИЯ ПО ДАННОМУ МОДУЛЮ (ТЕМЕ)</b> .....	2
1. Целеполагание по содержанию обучения. Отбор необходимого множества учебных элементов для включения в модуль (тему) .....	3
2. Целеполагание по уровню усвоения каждого учебного элемента .....	3
3. Разделение учебных элементов (с учетом требуемого уровня их усвоения) на два множества: проверяемых и непроверяемых существующими на сегодняшний день типами тестовых заданий .....	4
4. Разработка плана и спецификации теста .....	4
<b>РАЗДЕЛ 2. РАЗРАБОТКА ЗАДАНИЙ В ТЕСТОВОЙ ФОРМЕ</b> .....	5
5. Составление необходимого числа заданий в тестовой форме (одного или разного типа) для каждого из тестов-диагностируемых учебного элемента модуля с учетом требуемого уровня его усвоения .....	5
6. Составление несколько вариантов (одной сложности) для всех заданий, сконструированных на предыдущем этапе. Конструирование параллельных вариантов «теста» (пока на стадии «наборов заданий в тестовой форме») .....	12
7. Определение способа оценивания результатов тестирования студентов .....	12
<b>РАЗДЕЛ 3. ОТБОР, УТОЧНЕНИЕ И ОПРЕДЕЛЕНИЕ (НАЗНАЧЕНИЕ) ЗАДАНИЙ КАК ТЕСТОВЫХ</b> .....	13
8. Тестирование первой группы студентов с использованием одного или нескольких вариантов НЗТФ. Статистическая обработка результатов первого(ых) тестирования(ий). Отбор, уточнение и определение (назначение) заданий как тестовых .....	13
<b>РАЗДЕЛ 4. РАЗРАБОТКА ТЕСТА КАК СИСТЕМЫ ТЕСТОВЫХ ЗАДАНИЙ</b> .....	17
9. Конструирование, моделирование педагогического теста из набора тестовых заданий с опорой на критерии качества и эффективности .....	17
<b>ПРИЛОЖЕНИЯ</b> .....	18
Приложение 1. Рекомендации по технологии проектирования дидактических тестов .....	18
Приложение 2. Расчет коэффициента корреляции баллов испытуемых по заданию с баллами по всему тесту (корреляция задания с критерием) .....	19
Приложение 3. Словарь основных понятий, используемых в Руководстве .....	22
<b>ЛИТЕРАТУРА</b> .....	26
Список цитируемой литературы .....	26
Список рекомендуемой литературы .....	26

<sup>1</sup> В данном Руководстве не рассматриваются вопросы, связанные с внесением преподавателями своих педагогических тестов в оболочку «e-University» и разработкой конкретных сценариев тестирования. Эти темы изложены в справочном пособии к оболочке (Версия 4) — «10.Руководство по созданию вопросов» (<http://euniversity.bsu/eun4docs/10.pdf>). Также см. о сетевой образовательной платформе (СОП) eUniversity по адресу <http://intranet.bsu/eUniversity/>, где располагается **тренировочный модуль для преподавателей № 5 Создание тестов** (автор Анна Руденко) ([http://intranet.bsu/eUniversity/Трен.%20модуль%20№5\(создание%20теста,%20группы\)/untitled.htm](http://intranet.bsu/eUniversity/Трен.%20модуль%20№5(создание%20теста,%20группы)/untitled.htm))

<sup>2</sup> Автор с благодарностью примет критические замечания и рекомендации по совершенствованию данного Руководства. Адрес электронной почты: [krasnov@bsu.by](mailto:krasnov@bsu.by) (Краснову Юрию Эдуардовичу, нач. отдела стратегии образования Центра проблем развития образования Белгосуниверситета, г. Минск)

## **Введение. Этапы разработки педагогического теста: от заданий в тестовой форме через тестовые задания к тесту<sup>1</sup>**

Педагогические тесты проводятся с обучающее-тренинговой или контролирующе-оценивающей целью. Совмещение обеих задач мы получаем в случае самотестирования студентов. В обучающей функции допустимо использование разного рода *наборов заданий в тестовой форме* (не прошедших необходимую эмпирическую и статистическую проверку). Для промежуточного контроля необходимо использовать как минимум *системы из тестовых заданий* (прошедших необходимую эмпирическую и статистическую проверку). Для итогового контроля (в функции допуска к экзамену, зачета или итогового оценивания) необходимо использовать *тесты*. Определения базовых понятий: **задание в тестовой форме, тестовое задание, тест** — содержатся в Приложении 3 «Словарь основных понятий, используемых в Руководстве». Данным понятиям соответствуют три основных этапа разработки педагогического теста (включая подэтапы диагностического описания целей обучения, определения уровня усвоения каждого учебного элемента, разделения учебных элементов на проверяемые и непроверяемые существующими на сегодняшний день типами тестовых заданий). Текст Руководства будет разбит на 4 Раздела, но нумерацию параграфов при разработке теста мы оставили сквозной, чтобы подчеркнуть этим логику и последовательность шагов при разработке теста.

Задачи контролирующе-оценивающего тестирования органически вытекают из достоинств тестовой формы контроля — это достаточно непредвзятая и близкая к объективной оценка знаний группы студентов посредством экономичных, быстрых, технологичных процедур по стандартным формам. Ведущая идея традиционного теста — минимумом числом заданий, за короткое время, быстро, качественно и с наименьшими затратами сравнить знания как можно большего числа студентов.

## **Раздел I. Что надо сделать до разработки заданий в тестовой форме и построения теста? Диагностическое описание целей обучения по данному модулю (теме)<sup>2</sup>**

Выполнение требований данного раздела позволит определить, ЧТО подлежит тестированию, когда на следующем этапе (Раздел 2) будут разрабатываться задания в тестовой форме. Диагностическое описание целей обучения — это иной по сравнению с традиционными (аморфно, неопределенно и неинструментально) способ их описания, когда предполагается: 1) вполне определенное (конкретное) и однозначное описание целей обучения; 2) задание способов выявления, измерения и оценки степени их реализации. Цели формулируются через *результаты обучения, выраженные в действиях учащихся, которые могут быть однозначно и объективно оценены*. В качестве средства объективного контроля и выступают тесты. Тесты ориентированы на диагностические (т. е. поддающиеся

<sup>1</sup> Данное Руководство составлено по результатам работы и с учетом пожеланий участников общеуниверситетского семинара «Методика педагогических измерений. Организация тестирования», который работал на подготовительном этапе эксперимента по развитию самостоятельных форм учебной деятельности студентов (2005—2006 гг.) в рамках программы Белорусского государственного университета «Совершенствование организации, обеспечения и контроля качества самостоятельной работы студентов (2004—2009 гг.)». Автор благодарит З.П.Трофимову за ценные замечания и плодотворные дискуссии при обсуждении текста данного Руководства. Работая над его составлением, мы ориентировались на подход и работы российского тестолога В.С.Аванесова. В январе 2005 г. им был проведен недельный образовательный семинар с представителями *всех университетских кафедр* по освоению методологии и методики конструирования педагогических тестов.

<sup>2</sup> При написании Раздела I в значительной степени были использованы материалы пособия [Трофимова, 2005, 21—33].

диагностике, количественному измерению) параметры цели обучения и по определению должны измерять достижение именно этих целей. Если цель обучения назначена обычным образом (неопределенно и без указания параметров усвоения), например, через:

- 1) изучаемое содержание (например: «изучить содержание таких-то глав учебника»);
- 2) деятельность педагога (например: «ознакомить студентов с различными подходами к решению такой-то проблемы»);
- 3) деятельность студентов на занятии (например: «цель практического занятия — выполнение упражнений»);
- 4) внутренние процессы развития учащегося (например: «формировать интерес»), — то весьма затруднительно, а скорее невозможно, распознать ее реализацию в результатах обучения.

## 1. Целеполагание по содержанию обучения. Отбор необходимого множества учебных элементов для включения в модуль (тему)

**Модуль** — укрупненная единица обучения, включающая в себя несколько тем. Каждый из модулей характеризуется как единое целое — он имеет «свою» ключевую проблему, ведущую идею и цель изучения.

**Учебный элемент (УЭ)** — дидактическая единица, которая является оформленным под ту или иную образовательную задачу организованностью науки (явление, информация, гипотеза, доказательство, теорема, аксиома, эксперимент, факт, знание, теория, принцип, метод исследования и т.п.) или практики (умение, навык, компетенция, способ деятельности, ориентировочная основа деятельности и т.п.). Рекомендуется проектировать образовательный процесс с включением в УЭ различного рода «ключевых компетенций» как надпредметных знаний и умений.

Преподаватель отбирает или конструирует для конкретного модуля уникальный набор учебных элементов и представляет их не списком, а в виде **логической структуры модуля (ЛС)**, которая изображается в виде древовидной схемы, состоящей из нескольких этажей (порядков ЛС) узлов (УЭ), соединенных линиями. Большинство УЭ — это производные УЭ различного типа. Каждый УЭ выделяется на «своем» этаже по одному и тому же основанию.

Число, номенклатура (типы) и наименования всех УЭ по модулю должны быть известны студентам до начала прохождения модуля

## 2. Целеполагание по уровню усвоения каждого учебного элемента

Можно выделить 4 уровня усвоения каждого УЭ в зависимости от того как преподавателем (образовательным стандартом) планируется характер использования данного УЭ в научной и/или профессиональной деятельности<sup>1</sup>:

- 1). Две разновидности репродуктивной деятельности (**узнавание** и **воспроизведение**):
  - действие с подсказкой на *уровне узнавания* изученного УЭ (в задаче заданы цель, ситуация и действия по ее решению, от студента требуется дать заключение о соответствии всех трех компонентов);
  - действие по самостоятельному *воспроизведению* УЭ в *типовой* ситуации на основе *алгоритма* ориентировочной основы деятельности (ООД) (в задаче заданы цель и ситуация, а от студента требуется применить ранее усвоенные УЭ и действия по ее решению);
- 2). Две разновидности продуктивной деятельности (**эвристическая** и **творческая**):

<sup>1</sup> Для некоторых УЭ и видов деятельности принципиальным является задание *уровня автоматизации* исполняемых действий и операций.

- действие по самостоятельному *эвристическому* использованию УЭ в *нестандартной* ситуации на основе *трансформации* ранее известного *алгоритма* ориентировочной основы деятельности (ООД) (в задаче заданы цель, но не ясна ситуация, в которой цель может быть достигнута, а от студента требуется дополнить (уточнить) ситуацию и применить ранее усвоенные действия, или измененную, или «здесь и сейчас» разработанную ООД); в учебной ситуации студентом производится субъективно новая информация;
- действие по самостоятельному творческому использованию УЭ в новой ситуации на основе конструирования ранее неизвестного *алгоритма* ориентировочной основы деятельности (ООД) (в задаче известна лишь в общей форме цель деятельности, а поиску подвергается и подходящая ситуация и действия, ведущие к достижению цели, создается новая ООД); в учебной ситуации студентом производится объективно новая информация;

### **3. Разделение учебных элементов (с учетом требуемого уровня их усвоения) на два множества: проверяемых и непроверяемых существующими на сегодняшний день типами тестовых заданий**

Для УЭ (с учетом требуемого уровня их усвоения), которые не проверяются существующими на сегодняшний день типами тестовых заданий должны разрабатываться иные методы диагностики (письменная работа, эссе, собеседование, проект, комплексное квалификационное задание, задача и т. п.). Однако по возможности необходимо постараться разложить эти УЭ (сложные знания, умения, навыки, компетенции) на простые составляющие, усвоение которых может быть диагностируемо с помощью существующих типов тестовых заданий. Преподаватель в таком случае должен определить в итоговой (рейтинговой) оценке по модулю (предмету в целом) место и вес результатов **тестовых и нетестовых форм контроля** за учебными достижениями студентов.

Поскольку продуктивная творческая деятельность не содержит эталонного образца решения, постольку диагностика этого уровня усвоения практически невозможна в тестовой форме. По крайней мере можно предложить только следующее — всем студентам дается по стандартной форме проблема, не имеющая на сегодняшний день эталонного решения. Используется задание открытой формы, преподаватель анализирует предлагаемые решения на предмет корректного использования при этом тех или иных УЭ. Для удобства будем считать, что задания в тестовой форме не могут диагностировать 4 уровень усвоения УЭ, а также меру наличия у студентов «компетенций высшего уровня»<sup>1</sup>.

### **4. Разработка плана и спецификации теста**

Для набора всех тесто-диагностируемых УЭ по определенному модулю разрабатывается план их диагностики — **спецификация теста**, которая является примерной раскладкой процентного соотношения различных тем модуля (или дисциплины в целом) в содержании всего модуля (или дисциплины в целом) и схемой определения необходимого числа заданий по каждой теме содержания, исходя из важности темы и числа часов, отведенных на ее изучение в программе (относительный вес каждого УЭ, раздела темы и т.п.). Таким образом спецификация теста — это детальное определение структуры содержания модуля (дисциплины в целом) и состава всех УЭ, целей обучения в плане уровня усвоения каждого УЭ, количества тестовых заданий по каждому УЭ и тесту в целом

<sup>1</sup> Принципиальная критика возможностей сложившейся парадигмы тестирования для диагностики «компетенций высшего уровня» содержится в работах Дж. Равена [Равен, 1999; Равен, 2002].

(длина теста). В клетке таблицы — на пересечении горизонтальных строк (содержание) и вертикальных столбцов (целеобразование по уровню усвоения) — обозначается число намечаемых тестовых заданий и ставится прочерк если по УЭ, теме не планируется данный уровень усвоения (число заданий равно 0). Сокращенный вариант спецификации приведен в таблице 1<sup>1</sup>.

Таблица 1 Общая модель спецификации теста по модулю (данные приведены для примера)

Составные части (темы) изучаемого содержания модуля	Относительный вес каждой темы в % (учитывается кол-во часов и важность)	Учебные элементы, входящие в тему	Целеобразование по уровню усвоения каждого УЭ модуля				Всего (число тестовых заданий для каждого УЭ)
			Узнавание	Воспроизведение в знакомой ситуации	Применение в незнакомой ситуации	Исследовательская творческая деятельность	
Тема 1	30%	УЭ – 1.1	1	1		—	9
		УЭ – 1.2		2	2	—	
		УЭ – 1.3		1	2	—	
Тема 2	50%	УЭ – 2.1	1	1		—	15
		УЭ – 2.2		2	2	—	
		УЭ – 2.3		2	3	—	
		УЭ – 2.4		2	2	—	
Тема 3	20 %	УЭ – 3.1		1	2	—	6
		УЭ – 3.2		1	2	—	
Всего	100 %	9	2	13	15	0	30

## Раздел 2. Разработка заданий в тестовой форме<sup>2</sup>

### 5. Составление необходимого числа заданий в тестовой форме (одного или разного типа) для каждого из тестодиагностируемых учебного элемента модуля с учетом требуемого уровня его усвоения

После составления спецификации теста необходимо для каждого УЭ сконструировать задания в тестовой форме. На этом этапе от преподавателя требуется знание содержания предмета и типов заданий в тестовой форме, владение логикой и умением трансформировать фрагменты содержания учебной дисциплины в содержание заданий по определенным правилам и с учетом типа тестовой формы. Для каждого учебного модуля необходимо разработать порядка **100-150** «заданий в тестовой форме» для того, чтобы на следующем этапе можно было бы отобрать из них «тестовые задания» (около половины). При конструировании теста как эффективной системы из минимального числа заданий некоторые «тестовые задания» также будут отброшены. В результате остается несколько десятков заданий (более подробно см. параграф 9.1.).

<sup>1</sup> Таблица 1 составлена на основе соответствующей таблицы пособия [Трофимова, 2005, 32]. Конкретные данные приведены в качестве примера для лучшей наглядности. С более полными требованиями к спецификации теста можно ознакомиться по книге [Аванесов, 2002, 221–222].

<sup>2</sup> Данный раздел написан с опорой на работу В. С. Аванесова [Аванесов, 2005] (в некоторых местах мы использовали методические рекомендации главы 2.5 «Систематическое тестирование как средство управления и контроля самостоятельной работы» книги [Жук, 2005, 72–83]). Все примеры заданий в тестовой форме взяты из упомянутой работы В.С.Аванесова.

Расстановка заданий по **уровню трудности** осуществляется пока предварительно — с опорой на предполагаемое число и характер умственных операций, необходимых для успешного выполнения заданий студентами — **сложность задания** (впоследствии благодаря математико-статистическому анализу мера трудности каждого задания будет уточнена).

**Задание в тестовой форме** — вид *педагогического* задания, которое характеризуется стандартизированной, т. е. **тестовой**, формой фиксации, во-первых, указаний выполнить то или иное действие студенту и, во-вторых, результатов проделанной им работы (в узком смысле — ответов).

Наиболее распространенной является такой тип заданий (**закрытые**) в тестовой форме, когда варианты решения задания (результаты действия) предлагаются на выбор студенту. Среди них есть правильные(ый) и неправильные. Студент должен определиться в отношении предложенных ему вариантов. Иным типом являются **открытые задания**, в которых готовые решения (на выбор) не даются. Студент сам должен сформулировать решение (ответ) и зафиксировать его в специально отведенном месте.

### 5.1. Структура и элементы задания в тестовой форме

Каждый тип задания в тестовой форме имеет следующие универсальные элементы композиции:

1. **Инструкция** (например: «*Дополнить:* », набирается курсивом, в конце ставится двоеточие; при автоматизированном тестировании инструкцию лучше выводить на экран, вместе с каждым очередным заданием, особенно это полезно при использовании в одном тесте разных типов заданий)
2. **Содержание задания** (номер задания с точкой, содержание пишется ПРОПИСНЫМИ буквами, точка в конце не ставится; задание формулируется из точных терминов, без метафор, лишних слов и лишних знаков, как можно точнее, яснее и короче, полностью исключаются повторы слов, малопонятные, редко употребляемые, а также неизвестные студентам символы, иностранные слова; в хорошем задании от 7 до 10 слов; для достижения краткости лучше спросить о чем-нибудь одном; хорошо, когда задания содержат не более одного придаточного предложения; *стиль формулирования задания — не вопрос, а логика высказываний, позволяющая легко установить истинность/ложность формулировки*; не рекомендуется отрицательная формулировка заданий, или иначе, когда спрашивают о том, что не используется, не относится и т.д.).
3. **Ответы к заданиям** (номер со скобкой, текст ответов пишется *строчными* буквами, точка в конце не ставится; ответы желательно иметь короткими; задание должно иметь правильный ответ, не должно быть заданий, у которых нет решения, желательно чтобы все ложные ответы были **дистракторами** — правдоподобными, но неправильными ответами (в случае, когда трудно подобрать дистрактор, можно рекомендовать такой способ — на предварительном этапе задание дается студентам в открытой форме, типичные их ошибки становятся вариантами для выбора в заданиях закрытого типа); при этом абсурдных ответов быть не должно, в том числе таких ответов как: «правильного ответа нет», «все ответы правильные», «все ответы неправильные») (примеры оформления содержания задания и ответов к заданиям см. ниже)
4. **Оценка** (за правильный ответ). Как правило за правильный ответ дается 1 балл, за неправильный — ноль. Суммирование всех баллов, полученных испытуемым, дает число правильных ответов. Это число ассоциируется с уровнем его знаний и с

понятием «тестовый балл испытуемого». Есть и другие схемы оценивания. При этом может учитываться вес темы и/или УЭ в содержании модуля, а также в отдельных случаях может выставляться градуированная оценка в зависимости от числа правильно выбранных ответов и т.п.

*Часть заданий, которые используются для самообучения, самоконтроля студентами, сопровождаются указаниями к решению, а также точными и полными ответами.*

## **5.2. Типы заданий в тестовой форме («формы тестовых заданий» по В.С.Аванесову)**

Существует достаточно много типов заданий в тестовой форме. Мы рекомендуем опираться на классификацию В.С.Аванесова, который выделяет всего 4 известных на сегодня типа<sup>1</sup>. Каждый тип проверяет в той или иной степени разные виды знаний, поэтому их выбор зависит от учебного предмета, цели применения теста, способа обработки данных и, в целом, является творческим выбором самого преподавателя.

### **1 тип. Задания с выбором одного или нескольких правильных ответов.**

#### **1.1. Задания с выбором одного правильного ответа**

(типовые инструкции к каждому типу заданий даются курсивом).

*Обвести кружком номер правильного ответа:*

83. КИНЕТИЧЕСКАЯ ЭНЕРГИЯ ТЕЛА ЗАВИСИТ ОТ

- 1) массы
- 2) скорости
- 3) массы и скорости

Чем меньше число ответов к заданию, тем выше вероятность угадывания.

Рекомендуется иметь не менее **шести** ответов, там где это можно.

#### **1.2. Задания с выбором одного наиболее правильного ответа**

*Обведите кружком номер наиболее правильного ответа:*

---

<sup>1</sup> В данном тексте мы используем термин «тип» для классификации заданий в тестовой форме, относя понятие «тестовой формы» к стандартизированному краткому виду предъявления небольших заданий. Кроме того в лексике В.С.Аванесова невозможно сказать — форма задания в тестовой форме, а выражение «форма тестовых заданий», которое он использует, затрудняет понимание, потому что чаще всего относится к «заданиям в тестовой форме», а не к «тестовым заданиям», которые он строго различает.

При этом используемое некоторыми тестологами различение типов (закрытые, открытые) и видов (форм) тестовых заданий нам представляется несущественным, поскольку логически строгая классификация мало что дает преподавателю-практику, а утверждение, что уровни усвоения УЭ можно диагностировать строго определенными типами и/или видами тестовых заданий [Беспалько, 1988, 1995; Трофимова, 2005, 36—40] нам представляется неубедительным. Обращаем внимание читателя на тот факт, что З. А. Трофимова, вслед за В. П. Беспалько, в своем пособии рекомендует использовать для диагностики усвоения УЭ на 3 и 4 уровне такие «виды тестовых заданий», которые неизвестны науке и практике тестологии, т. е. они еще требуют своей разработки («нетиповая задача», «проблемный вопрос», «на постановку проблемы», «на поиск путей решения проблемы», «на проверку правильности ее решения»).

Нам представляется, что достижение данной диагностической задачи может быть в *некоторых частных случаях* выполнено и с использованием уже известных «форм тестовых заданий» (по В.С.Аванесову). Все зависит от того, удастся ли преподавателю-разработчику смоделировать эвристическую и/или творческую мыследеятельность студента, представить ее в виде последовательности действий. Заложить эту интеллектуальную модель в форму тестового задания (практически любую) останется делом техники. Проблемой, однако, является, по нашему мнению, как раз неразработанность для большинства случаев соответствующих типов «заданий в тестовой форме».

## 16. {счетные, несчетные} МНОЖЕСТВА ОБРАЗУЮТ ЧИСЛА

- 1) целые
- 2) натуральные
- 3) рациональные
- 4) иррациональные
- 5) целые и натуральные
- 6) натуральные и рациональные
- 7) целые, натуральные и рациональные
- 8) целые, натуральные, иррациональные
- 9) целые, натуральные, рациональные и иррациональные

В такого рода задании желательно, используя большой выбор, давать градуированную оценку: 0, 1 и 2 — за выбор частично правильного, чуть более правильного и абсолютно правильного, хотя возможна и двухбалльная оценка (0 и 1). Сложность оценивания ответов таких заданий связана с тем, что все ответы частично правильны. Поэтому с методикой оценивания таких заданий желательно ознакомиться по литературе.

## 1.3. Задания с выбором нескольких правильных ответов

*Вашему вниманию предлагаются задания, в которых могут быть один, два, три и большее число правильных ответов. Обведите кружком номера всех правильных ответов: (в компьютерном варианте концовка инструкции такова: «...Отвечая на задания, нажимайте на клавиши с номерами всех правильных ответов»)*

## 2. ФИЛОСОФСКИЕ ПОНЯТИЯ

- |               |                |
|---------------|----------------|
| 1) бытие      | 6) знание      |
| 2) сознание   | 7) свобода     |
| 3) развитие   | 8) качество    |
| 4) культура   | 9) революция   |
| 5) диалектика | 10) количество |

Рекомендуется использовать задания с выбором нескольких правильных ответов. Студент должен определить число, а затем указать конкретно эти правильные ответы. Это снижает на порядок вероятность угадывания правильных ответов. Общее число ответов желательно иметь в диапазоне от 5 до 12. Четное число смотрится эстетичнее. Число правильных ответов в разных заданиях различно, поэтому такие задания труднее, они проверяют знания полнее, глубже и точнее, чем задания с выбором одного правильного ответа. Нежелательно делать все ответы правильными, чтобы не подталкивать студентов к искусственным ошибкам. Можно порекомендовать такую шкалу оценивания: за полностью правильное решение дать 3 балла, за одну допущенную ошибку снимать один балл, за вторую ошибку — еще один балл, за третью — снимать последний балл. За последующие ошибки баллы лучше не снимать. Более точно оценки за разные варианты выполнения можно найти посредством шкалирования заданий по уровню их трудности.

**2 тип. Задания открытой формы**

*Дополнить:*

1. ПЕРВЫМ ГРЕЧЕСКИМ ФИЛОСОФОМ СЧИТАЕТСЯ \_\_\_\_\_

В заданиях открытой формы нет готовых ответов. Соответственно, нечего и выбирать. Студент сам пишет ответ в бланке, или набирает его с клавиатуры компьютера. При разработке таких заданий нужно стремиться к тому, чтобы: содержание заданий было понятно всем, без исключений, испытываемым; все утверждения делались бы, по возможности,



короткими; добавлялось бы как можно меньше слов или знаков и желательно что-то одно; добавляемое слово, для лучшего понимания, желательно ставилось бы ближе в конец задания; не делать задания с двумя, тремя пропусками. Ответом является дополняемое слово, знак, формула и т.п. Обычно используется двухбалльная шкала для оценивания ответов (0 и 1).

### 3 тип. Задания на установление соответствия

Установить соответствие:

17. ФОРМУЛА

ИСПОЛЬЗУЕТСЯ ДЛЯ  
ВЫЧИСЛЕНИЯ

- |                           |                                |
|---------------------------|--------------------------------|
| 1. Лагранжа               | А) несобственного интеграла    |
| 2. Лопиталья              | Б) неопределенного интеграла   |
| 3. Трапеций               | В) определенного интеграла     |
| 4. Половинного<br>деления | Г) пределов                    |
|                           | Д) производной в средней точке |
|                           | Е) середины отрезка            |
|                           | Ж) корней уравнений            |

Ответы: 1 \_\_\_\_, 2 \_\_\_\_, 3 \_\_\_\_.

В заданиях этого типа студент должен установить соответствие элементов одного столбца элементам другого столбца. Этот тип заданий хорошо проверяет знания о взаимосвязи определений и фактов, авторов и их произведений, форм и содержания, сущности и явлений, о соотношении между различными предметами, свойствами, законами, формулами, датами. Помимо уже упомянутых, данный тип заданий содержит такие элементы композиции как: *названия двух столбцов* — заголовки (должны относиться ко всем элементам представляемого класса); *элементы этих столбцов* (в правом столбце лучше располагать избыточные элементы — дистракторы, число элементов правого столбца примерно в два раза больше числа элементов левого столбца, чтобы последняя пара не подбиралась методом исключения); *номера и буквы* (идентификаторы элементов, цифры лучше писать в левом столбце, буквы — в правом, номера отделяются от содержания столбцов точкой, а буквы — круглой скобкой и пробелом); *строка ответов* (пишется точно посередине задания, иногда на свободное место в первом столбце, в конце строки лучше ставить точку).

### 4 тип. Задания на установление правильной последовательности

Установить правильную последовательность:

16. РЕШЕНИЕ ЗАДАЧ ПО ДИНАМИКЕ

1. СОБЫТИЯ ФЕВРАЛЯ-ОКТЯБРЯ 1917 г.

- |  |   |
|--|---|
| <input type="checkbox"/> – выбрать систему отсчета                   | <input type="checkbox"/> – VI съезд РСДРП (б)             |
| <input type="checkbox"/> – записать 2 закон Ньютона                  | <input type="checkbox"/> – отречение царя Николая II      |
| <input type="checkbox"/> – записать в скалярном виде 2 закон Ньютона | <input type="checkbox"/> – приезд Ленина                  |
| <input type="checkbox"/> – решить систему                            | <input type="checkbox"/> – создание Петроградского совета |
| <input type="checkbox"/> – расставить силы                           | <input type="checkbox"/> – взятие Зимнего дворца          |
| <input type="checkbox"/> – проанализировать ответ                    | <input type="checkbox"/> – Корниловский мятеж             |
| <input type="checkbox"/> – получить систему уравнений                | <input type="checkbox"/> – ликвидация двоевластия         |
| <input type="checkbox"/> – сделать схематический рисунок             | <input type="checkbox"/> – II съезд Советов               |

22. СЮЖЕТ «КАПИТАНСКОЙ ДОЧКИ»  
ПУШКИНА

30. ТЕРМОХИМИЯ

- |   |                                    |
|---|------------------------------------|
| <input type="checkbox"/> – решение отца Гринева | <input type="checkbox"/> – наука   |
| <input type="checkbox"/> – Маша спасает Гриневу | <input type="checkbox"/> – раздел  |
|   | <input type="checkbox"/> – энергия |

- |   |  |
|---|--|
| <input type="checkbox"/> – Пугачевское восстание        | <input type="checkbox"/> – тепловая    |
| <input type="checkbox"/> – по дороге на место службы    | <input type="checkbox"/> – взаимное    |
| <input type="checkbox"/> – арест Петруши Гринева        | <input type="checkbox"/> – химическая  |
| <input type="checkbox"/> – встречи в заснеженной степи  | <input type="checkbox"/> – изучающая   |
| <input type="checkbox"/> – жизнь в Белогорской крепости | <input type="checkbox"/> – превращение |

Данный тип заданий позволяет проверять культуру алгоритмического мышления (нацеленной на определение наилучшей последовательности действий при решении тех или иных задач), знания, умения, навыки по установлению правильной последовательности различных по содержанию «объектов»: 1) исторических событий; 2) технологических действий, операций, расчетов (в том числе при решении типовых задач); 3) процессов; 4) терминов в научных определениях; 5) событий сюжетов различного рода художественных произведений и т.п. *Название задания* фиксирует то, что должен продемонстрировать студент, ключевое слово названия лучше всего писать в именительном падеже. *Содержание задания* включает в себя название и ранжируемые «объекты». *Место для ответов* — это прямоугольники, нарисованные слева, против названия каждого элемента. В них нужно проставлять соответствующие цифры (ранги). В самом задании элементы ставятся в случайном порядке, чтобы в их расположении не было никакого намека на правильный порядок. Для предотвращения угадывания окончания всех слов лучше писать в именительном падеже. Оценка за выполнение каждого задания может варьироваться в зависимости от его важности и трудности. Чаще всего, однако, используется дихотомическая оценка 1/0, при которой один балл дается за правильную расстановку всех рангов в задании, ноль — при ошибке в ответе. Возможно использовать и такое правило: безошибочное решение оценивается, например, 3 баллами; ошибка конце задания — 2 баллами; ошибка в середине — 1 баллом; ошибка в начале — 0 баллов.

### 5.3. Требования к заданиям в тестовой форме

В отличие от распространенной вопросно-ответной формы заданий, в качестве новой основы используется **логика высказываний**. Преимущества этой основы вытекают из отличия высказываний от вопросов: истинность или ложность высказывания легко определяется по логическим правилам, в то время как вопросы сами по себе ни истинны, ни ложны. Кроме этого формулировки вопросов занимают больше места нежели утверждения (см. Приложение 3, определение термина «**вопрос**»).

Задание в тестовой форме (в отличие от нетестовой) отвечает требованиям:

- **цель** (для: слабых или сильных студентов; аттестации или профессионального отбора; обучения или контроля; диагностики того или иного вида знаний, что влияет на правильность выбора типа задания и т.п.)
- **краткость** (см. 5.1. пункт 2. «Содержание задания»; краткость является следствием преимуществ логической формы высказывания)
- **технологичность** (возможность использования в системах компьютерного обучения и контроля, этому способствует ясность формы и содержания задания, дихотомическая система оценивания 1/0 и т.п.)
- **соответствие определенному типу** (недопустимо смешивать типы в одном задании)
- **логическая форма высказывания**
- **наличие определенного места для ответов**
- **одинаковость правил оценки ответов** (ни одному испытуемому не дается никаких преимуществ перед другими, все отвечают на одни и те же задания, всем дается одинаковое время, правила оценки определяются заранее и абсолютно одинаково применяются ко всем испытуемым)

- **правильность расположения элементов задания** (является требованием, помогающим испытуемому не тратить время на определение места для ответов и быстрее фиксировать свои решения)
- **одинаковость инструкции для всех испытуемых** (это условие справедливости и стандартизации педагогических измерений)
- **адекватность инструкции форме и содержанию задания** (взаимное соответствие всех этих трех компонентов, что необходимо для выполнения заданием своей функции; несоответствие формы содержанию и, наоборот, вызывает ошибку понимания смысла задания и может приводить к неправильному ответу).

#### ***5.4. Возможности ситуационных заданий в тестовой форме для проверки комплексных компетенций и практических умений действовать в ситуациях***

**Ситуационные задания** относятся к так называемым «**системам заданий в тестовой форме**» (СЗТФ), которые надо отличать от промежуточной стадии при разработке теста — «**набора заданий в тестовой форме**» (НЗТФ). И так *СЗТФ образуют такие задания, которые имеют все признаки, присущие заданиям в тестовой форме, но не обладают свойствами тестовых заданий, а поэтому они в принципе тестом не являются.* СЗТФ применяют для педагогического оценивания, не претендующего на соответствие строгим требованиям научного измерения, что свойственно тестам. При применении СЗТФ получаемые студентами баллы сводятся к обычным оценкам (в результате субъективного оценивания баллов), которые не могут интерпретироваться как результат педагогического измерения.

К СЗТФ относятся **цепные** (правильный ответ на последующее задание зависит от ответа на предыдущее задание), **тематические** (совокупность заданий любой формы, созданная для контроля по одной изученной теме), **текстовые** (совокупность заданий открытой формы, созданных для автоматизированного контроля знаний учебного текста), **ситуационные** (педагогически переработанный фрагмент профессиональной деятельности специалиста; разрабатываются для проверки знаний и умений испытуемых действовать в практических, ситуациях, включая экстремальные; дается текстовой описание ситуации и к ней прикладывается система заданий в тестовой форме). ***В СЗТФ реализуется принцип — вероятность правильного ответа на последующее задание может зависеть от вероятности правильного ответа на предыдущие задания. Это нарушает принцип статистической независимости тестовых заданий (аксиома локальной независимости), который положен в основу создания теста как формальной системы. Поэтому задания в таких системах (СЗТФ) теста не образуют.***

Однако можно всячески рекомендовать использование, в особенности, ситуационных заданий для организации СРС (когда к ситуации даются обоснования правильных ответов, что позволяет студентам формировать аргументированные и практико-ориентированные знания) и проведения итоговой аттестации выпускников образовательных учреждений, направленной на диагностику интегративной компетентности студентов по применению знаний в ситуации целостной профессиональной деятельности <sup>1</sup>. Ситуационные задания

<sup>1</sup> Рекомендуется следующая стратегия использования ситуационных заданий для проведения итоговой аттестации выпускников вуза, которая апробирована на кафедре терапевтической стоматологии Московского медицинского стоматологического университета (утверждено Министерством здравоохранения РФ, содержание заданий разработано совместно с В. С. Аванесовым, форма заданий не имеет аналогов в мире). Та или иная ситуация выбирается (по аналогии с экзаменационным билетом) самим выпускником, что создает предпосылку для объективированной итоговой оценки. Испытуемый отвечает на СЗТФ, приложенную к ситуации, после чего идет на собеседование. Предмет собеседования — действия врача в лечебной ситуации, по которой он выполнил задания в тестовой форме; возможные ошибки, причины их

могут входить в «фонд комплексных квалификационных заданий» (по В. П. Беспалько), направленных на контроль требуемого уровня подготовленности специалиста.

## **6. Составление несколько вариантов (одной сложности) для всех заданий, сконструированных на предыдущем этапе.**

### **Конструирование параллельных вариантов «теста» (пока на стадии «наборов заданий в тестовой форме»)**

На предыдущем этапе для каждого учебного модуля необходимо было разработать порядка **100-150** «заданий в тестовой форме» для того, чтобы на следующем этапе можно было бы отобрать из них «тестовые задания» (около половины). Полученный НЗТФ должен быть «размножен» до нескольких вариантов (В.С.Аванесов рекомендует иметь до 15 вариантов). Во-первых, это облегчает тестирование больших групп студентов (затрудняется списывание, угадывание и т.п.), во-вторых, **надежность тестовых результатов** (в практическом смысле надежность обычно понимается как мера одинаковости, повторяемости или связанности двух измерений одного и того же качества, одним и тем же тестом, или его параллельными вариантами) часто выражается мерой корреляции между двумя **параллельными вариантами теста**.

Поэтому в идеале надо стремиться сделать несколько вариантов теста, а для этого нужно разрабатываться варианты одного и того же задания, одинаковые по содержанию, типу, трудности. При составлении *параллельных заданий* теста желательно использовать фасет — форму записи нескольких вариантов одного и того же задания. Принцип фасетности содержания — это главный принцип создания параллельных тестов.

Пример:

88. ПРИ НАМОКАНИИ {хлопкового, вискозного...} ВОЛОКНА ЕГО ПРОЧНОСТЬ
- 1) увеличивается
  - 2) остается без изменений
  - 3) уменьшается

Составленные параллельные НЗТФ желательно обсудить с экспертами или коллегами.

## **7. Определение способа оценивания результатов тестирования студентов**

(за сколько правильно выполненных заданий студент получает ту или иную оценку или, например, допуск к экзамену)

Важно понимать — не может существовать универсальный принцип оценивания результатов тестирования, поскольку интерпретация результатов зависит от цели тестирования, содержания теста, уровня подготовленности отвечающих, типа интерпретации (предметно-педагогическая, критериально-ориентированная, нормативно-ориентированная). Тест который решают большинство студентов возможно и хорош для проверки по критерию минимального образовательного стандарта, но скорее всего как таковым тестом не является (точка зрения В.С.Аванесова), поскольку нет достаточной **дисперсии** результатов. Тест,

предназначенный для отбора из числа сильных студентов самых лучших, покажет полное отсутствие знаний даже на группе средних студентов.

Поэтому при определении способа оценивания результатов тестирования студентов преподаватель должен определиться самостоятельно, ответив на следующие вопросы:

- сколько баллов считать незачетным? \_\_\_\_\_
- как переводить тестовые баллы в отметку? \_\_\_\_\_
- допустимое округление и в какую сторону? \_\_\_\_\_
- критерии перевода результатов тестирования в 10-бальную шкалу. \_\_\_\_\_

Однако в случае, когда осуществляется стандартизация теста, по итогам которой результаты респондентов будут иметь *нормальное распределение* (меньше слабых, больше всего средних, меньше сильных), можно опираться на рекомендацию В.С.Аванесова — *минимально допустимый коэффициент выполнения теста* (К) должен быть 0.7 (отношение числа выполненных заданий к общему числу заданий в тесте).

## **Раздел 3. Отбор, уточнение и определение (назначение) заданий как тестовых<sup>1</sup>**

### **8. Тестирование первой группы студентов с использованием одного или нескольких вариантов НЗТФ. Статистическая обработка результатов первого(ых) тестирования(ий). Отбор, уточнение и определение (назначение) заданий как тестовых**

Теоретически, до тестирования первых студентов необходимо провести пробное тестирование на иной выборке с целью отбраковки плохо работающих заданий и дистракторов. По рекомендации специалистов такая проверка может быть осуществлена на типичной выборке в 100 чел. (минимум). Однако в условиях учебного процесса скорее всего придется эмпирическую проверку и опробывание НЗТФ осуществлять по ходу обучения, делая это на первой группе студентов. Данный этап (его можно назвать условно «апробацией теста») является обязательным при составлении тестовых заданий и конструировании тестов. Об удаче или неудаче каждого заранее составленного (умозрительно) задания могут определенно свидетельствовать только результаты эмпирического исследования. Только после экспериментальной проверки тестовых результатов и их статистической обработки делается вывод о валидности (пригодности, действенности) и надежности результатов теста. Следует говорить о **валидности и надежности тестовых результатов**, а не теста как такового (это обосновывается в работах В.С.Аванесова).

Превращение заданий в тестовой форме в тестовые задания начинается с момента статистической проверки каждого задания на наличие у них тест-образующих свойств. Каждое задание, прежде чем стать тестовым, должно пройти этап статистического и тестологического обоснования.

#### ***8.1. Отбор по математико-статистическим требованиям работающих заданий и отброс неработающих***

---

<sup>1</sup> Данный раздел написан с опорой на работу [Аванесов, 2002] (в некоторых местах использовался материал лекций В. С. Аванесова для преподавателей БГУ в январе 2005 г.)

**Тестовое задание** — это *составная единица* теста, отвечающая всем требованиям к заданиям в тестовой форме (технологичности, чистоты формы и предметного содержания) и, кроме того, *статистическим требованиям*:

- 1) известной (меры) трудности;
- 2) дифференцирующей способности (достаточной вариации тестовых баллов);
- 3) положительной корреляции баллов испытуемых по заданию с баллами по всему тесту
- 4) а также другим математико-статистическим требованиям, которые в данной инструкции опускаются («подходящий геометрический образ задания» и т.д.).

Из перечисленных требований вытекает обязательность эмпирической проверки заданий на выборочной совокупности испытуемых и применения статистических методов обработки данных. Эта работа ведется посредством **матриц тестовых результатов** (которые становятся таблицами, когда добавляются результаты их обсчета).

Таблица 2 Тестовых результатов (пример взят из работ В.С. Аванесова)

№№	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>	Y <sub>i</sub>	p <sub>i</sub>	q <sub>i</sub>	p <sub>i</sub> /q <sub>i</sub>	lnp <sub>i</sub> /q <sub>i</sub>
1.	1	1	1	0	1	1	1	1	1	1	9	.90	.10	9	2.20
2.	1	1	0	1	1	1	1	1	1	0	8	.80	.20	4	1.39
3.	1	1	1	1	0	1	1	0	1	0	7	.70	.30	2.33	.85
4.	1	1	1	1	0	1	0	1	0	0	6	.60	.40	1.50	.40
5.	1	1	1	1	1	1	0	0	0	0	6	.60	.40	1.50	.40
6.	1	1	1	1	0	0	1	0	0	0	5	.50	.50	1.00	0
7.	1	1	0	1	1	0	1	0	0	0	5	.50	.50	1.00	0
8.	1	1	1	1	1	0	0	0	0	0	5	.50	.50	1.00	0
9.	1	0	1	0	1	1	0	0	0	0	4	.40	.60	.66	-.42
10.	0	1	1	0	0	0	0	1	0	1	4	.40	.60	.66	-.42
11.	1	1	1	0	0	0	0	0	0	0	3	.30	.70	.43	-.84
12.	1	1	0	0	0	0	0	0	0	0	2	.20	.80	.25	-1.39
13.	1	0	0	0	0	0	0	0	0	0	1	.10	.90	.11	-2.21
R <sub>j</sub>	12	11	9	7	6	6	5	4	3	2	65				
W <sub>j</sub>	1	2	4	6	7	7	8	9	10	11					
p <sub>j</sub>	.923	.846	.692	.538	.462	.462	.385	.308	.231	.154	5				
q <sub>j</sub>	.077	.154	.308	.462	.538	.538	.615	.692	.769	.846					
p <sub>j</sub> q <sub>j</sub>	.071	.130	.213	.248	.248	.248	.236	.213	.178	.130					
q <sub>j</sub> /p <sub>j</sub>	.083	.182	.445	.859	1.164	1.164	1.597	2.246	3.329	5.493					
lnq <sub>j</sub> /p <sub>j</sub>	-2.489	-1.704	-.810	-.152	.152	.152	.468	.809	1.202	1.703					

Строки (от 1 до 13) представляют испытуемых. Столбцы (от 1 до 10) — задания. На пересечении записывается исход противоборства каждого испытуемого (i) с предлагаемым ему заданиями (j). В целях удобства обсчета и применения наиболее простых математических моделей рекомендуется допускать только два исхода (правильное решение — 1 балл, ошибочное — 0). В матрице проводится два упорядочения. Одно касается испытуемых. В первой строке фиксируются баллы самого успешного испытуемого, во второй менее, и т.д. В первом столбце располагается самое легкое задание (по нему имеется наибольшее число правильных ответов), на втором — менее легкое и т.д.

Для проверки тестовых свойств заданий, с данными таблицы выполняется ряд расчетов. Сложение по строке дает **тестовый балл каждого испытуемого** (Y<sub>i</sub>). Сложение по столбцам — **количество правильных ответов в каждом задании** (R<sub>j</sub>). Смысл переменных i и j — любой(ое) испытуемый/задание. Символ W<sub>j</sub> означает **число неправильных ответов испытуемых в каждом задании**. Поскольку показатели R<sub>j</sub> и W<sub>j</sub> зависят от числа испытуемых (N), для получения сопоставительных характеристик они делятся на N. В результате получается статистика p<sub>j</sub> (p<sub>j</sub> = R<sub>j</sub> / N) — **доля правильных ответов на задание j**, а также статистика q<sub>j</sub> (q<sub>j</sub> = W<sub>j</sub> / N) — **доля неправильных ответов на задание j**. В статистике естественным образом принимается p<sub>j</sub> + q<sub>j</sub> = 1.

**Суммирование баллов всех испытуемых** фиксируется (в данном случае оно равно 65). **Средний арифметический тестовый балл** в данной группе испытуемых —  $M = 65/13 = 5$ . Он равен сумме всех  $p_j$  ( $\sum p_j = M$ ) потому, что  $p_j$  является также и средним арифметическим для каждого задания.

Для заданий, в которых используется только дихотомическая оценка (1 или 0) дисперсия (стандартная мера вариации баллов) определяется по сравнительно простой формуле:  $s_j^2 = p_j q_j$ . Эти значения дисперсий (рассчитанные по этой формуле) представлены в пятой строке нижней части таблицы 2.

Показатели  $\ln q_j/p_j$  (логит трудности задания) и  $\ln p_i/q_i$  (логит уровня знаний) мы в данном тесте не рассматриваем (более подробно см. работы В.С. Аванесова).

Итак, рассмотрим обязательные тестологические, статистические характеристики заданий, которые необходимо просчитать, чтобы определяться относительно включения данного задания в тест.

**8.1.1. Трудность (мера) тестового задания.** Поскольку тест является системой заданий возрастающей трудности, постольку необходимо знать меру трудности каждого задания. Разумно расположить задания по мере возрастания их трудности, чтобы каждый студент в рамках отведенного времени мог дойти и успешно преодолеть только посильные для себя задания, не тратя время на решение непосильных для себя заданий. В теории тестов долгое время мерой трудности был показатель  $p_j$ , сейчас используется  $q_j$  — поскольку с его ростом растет и трудность задания. Логит ( $\ln q_j/p_j$ ) — сравнительно новая единица измерения трудности задания.

В классической теории тестов многие годы рассматривались только эмпирические показатели трудности. В новых вариантах психологических и педагогических теорий тестов больше внимание стало уделяться характеру умственной деятельности учащихся в процессе выполнения тестовых заданий различных форм.

**8.1.2. Дифференцирующая способность задания (достаточная вариация тестовых баллов).** Отсутствие вариации свидетельствует либо об одинаковости испытуемых, либо о несостоятельности инструмента и процедуры оценивания, но скорее о втором. Задания по которым нет вариации тестовых баллов (очень легкие, которые все решают, и очень трудные, которые никто не решает) в тест не включаются.

Эти показатели позволяют оценить меру рассеяния данных относительно средней арифметической

Стандартной мерой вариации являются значения **дисперсии баллов** ( $s_j^2$ ), а также корень квадратный из этого значения, называемый **стандартным отклонением** ( $s_j$ ). Для заданий, в которых используется только дихотомическая оценка (1 или 0) дисперсия определяется по сравнительно простой формуле:  $s_j^2 = p_j q_j$  (где  $p_j$  и  $q_j$  - доли правильных и неправильных ответов в каждом задании).

Для данных, представленных другими числами, дисперсия находится в два приема. Вначале находят **сумму квадратов отклонений тестовых баллов от среднего арифметического балла**:  $SS_y = \sum (X - M)^2$ . У этого важного статистического показателя есть существенный недостаток, заключающийся в зависимости от числа испытуемых: при прочих равных условиях, чем больше группа, тем большей оказывается  $\sum (X - M)^2$ , что делает этот показатель несопоставимым для групп с разным числом испытуемых. Поэтому для исправления отмеченного недостатка используют второй прием - делят  $SS_y$  на  $N - 1$ , где  $N$  - число испытуемых в группе. В результате получается стандартный показатель вариации тестовых баллов, называемый дисперсией ( $s_y^2$ ), или по-старому, **вариансой**<sup>1</sup>.

Стандартное отклонение  $s_y$  является общепринятой мерой вариации тестовых баллов. Вычисляется по формуле (корень квадратный из значения дисперсии  $s_y^2$ ):

$$s_y = \sqrt{\frac{SS_y}{N - 1}}$$

<sup>1</sup> От англ. Variance - изменчивость

**8.1.3. Положительная корреляция баллов испытуемых по заданию с баллами по всему тесту (корреляция задания с критерием).** *Корреляция задания с критерием* ( $r_{xy}$ ) является статистическим требованием к тестовым заданиям. Это распространенное в практике название, представляет собой сокращенный вариант более правильного понятия - корреляция оценок, полученных испытуемыми в задании, с оценками, полученными ими же по какому-либо критерию, например, по сумме баллов. Корреляция является стандартной мерой дифференцирующей способности задания. Чем выше значение корреляции задания с суммой баллов, тем выше его дифференцирующая способность. При проверке тестовых свойств заданий в качестве критерия, для начала, используется сумма баллов испытуемых, полученная ими по всем заданиям пробного варианта теста. Задание в тестовой форме нельзя называть тестовым, если оно не коррелирует (не связано) с отмеченной суммой.

Расчет коэффициента корреляции баллов испытуемых по заданию с баллами по всему тесту (корреляция задания с критерием) см. в Приложении 2.

Анализу тестовых свойств заданий очень способствует расчет **полной корреляционной матрицы** (Таблица 3), в которой представляют корреляции каждого задания со всеми остальными заданиями, а также с суммой баллов (предпоследний столбец). Расчет такой матрицы можно делать с опорой на стандартные статистические программы: Statistica, SPSS и др.

Таблица 3 Полная корреляционная матрица (пример взят из работ В.С. Аванесова)

Табл.10.3

ЗАДАНИЯ	1	2	3	4	5	6	7	8	9	10	$r_{jy}$	$r_{jy}^2$
1	1.000	0.123	-0.192	0.312	0.267	0.267	0.228	-0.433	0.158	-0.677	0.132	0.017
2	0.123	1.000	0.178	0.460	-0.033	-0.033	0.337	0.284	0.233	0.182	0.488	0.238
3	-0.192	0.178	1.000	0.051	-0.051	0.283	-0.158	0.083	-0.030	0.284	0.305	0.093
4	0.312	0.460	0.051	1.000	0.238	0.238	0.415	-0.051	0.141	-0.460	0.494	0.244
5	0.267	-0.033	-0.051	0.238	1.000	0.381	0.220	0.051	0.225	0.033	0.494	0.244
6	0.267	-0.033	0.283	0.238	0.381	1.000	0.220	0.386	0.592	0.033	0.706	0.498
7	0.228	0.337	-0.158	0.415	0.219	0.219	1.000	0.158	0.693	0.101	0.652	0.424
8	-0.433	0.284	0.083	-0.051	0.051	0.386	0.158	1.000	0.426	0.639	0.534	0.285
9	0.158	0.233	-0.030	0.141	0.225	0.592	0.693	0.426	1.000	0.272	0.752	0.565
10	-0.677	0.182	0.284	-0.460	0.033	0.033	0.101	0.639	0.272	1.000	0.293	0.086

Чем выше значения  $r_{xy}$ , тем больше вероятность превращения задания в тестовой форме в тестовое задание, то есть быть включенным в тест. Особенно заметно вероятность повышается при  $r > 0,400$ . Если взять  $r^2 \times 100\%$ , то получим значение так называемого коэффициента детерминации, выраженного в удобной, для интерпретации, процентной мере связи задания с суммой баллов. Для взятого примера коэффициент детерминации у седьмого задания равен  $(0,652)^2 \times 100\% = 42,5\%$ , что можно интерпретировать так: 42,5% вариации суммы тестовых баллов испытуемых по всем заданиям связано с вариацией баллов по одному только седьмому заданию, что указывает на очень высокий потенциальный вклад седьмого задания в общую дисперсию тестовых баллов.

При значениях  $0,2 < r_{xy} < 0,5$  во внимание начинают приниматься и другие характеристики, такие как мера трудности заданий, мера корреляции задания с другими заданиями, факторная чистота задания и другие, рассмотрение которых потребовало бы здесь другого стиля изложения. В качестве нижней границы включения заданий в тест можно рассматривать значение  $r_{xy} = 0,2$ . Нулевая корреляция свидетельствует об отсутствии у задания системных свойств, присущих для тестового задания. Такие задания, равно как и задания с отрицательными значениями  $r_{xy}$ , устраняются из тестовых материалов, как не выдержавшие эмпирической проверки.



## **8.2. Отбор в каждом тестовом задании работающих дистракторов и отбраковка неработающих**

Для отбраковки неработающих дистракторов нужно посчитать: сколько % выбирают 1-й дистрактор, 2-ой и т.д. В тестовом задании не должно быть ни одного не работающего дистрактора (ответа, который никто не выбирает). Все дистракторы всех заданий должны пройти проверку эмпирическим путем. Если ответ никто не выбирает, его просто вычеркивают и ищут замену.

## **8.3. Составление набора тестовых заданий для конструирования теста**

Шансы попасть в тест имеют только те задания в тестовой форме, у которых правильная форма, точно сформулированное и корректно отобранное содержание, а также есть математико-статистические характеристики, необходимые для создания теста как педагогической системы. Из этого вытекает обязательность эмпирической проверки свойств заданий на выборочной совокупности испытуемых, а также необходимость применения статистических методов обработки данных. Обязательность вытекает из свойства тестовых заданий быть частью метода педагогических измерений. На данном этапе происходит отбор хорошо работающих и отбраковка неработающих заданий в тестовой форме и образование, конструирование «тестовых заданий», которые в принципе могут попасть в тест. Однако сама процедура моделирования теста предполагает выбор уже из числа тестовых заданий.

# **Раздел 4. Разработка теста как системы тестовых заданий**

## **9. Конструирование, моделирование педагогического теста из набора тестовых заданий с опорой на критерии качества и эффективности**

Наличие достаточного числа тестовых заданий позволяет перейти к разработке *теста как системы*, обладающей целостностью, составом и структурой.

Тест представляет собой систему заданий, каждое из которых необходимо для выполнения тестом своей функции. Удаление хотя бы одного из системного задания приводит к возникновению пробела на континууме проверяемых знаний, и поэтому снижает качество измерения. Тест — это шкала, а не «проба», «проверка». Для повышения точности измерения применяется расстановка тестовых заданий в порядке возрастающей сложности. В тест включается минимальное число заданий необходимое для получения точных результатов. Тестирование без указания предполагаемой интерпретации результатов не имеет научного смысла.

К основным статистическим характеристикам педагогического теста относятся:

- **средние арифметические** (M);
- **показатели вариации**;
- **показатели связи**;
- **гистограмма** (показывающая частоту встречаемости тех или иных тестовых баллов и меру их отклонения от нормального распределения; более подробно см.: [Аванесов, 2002, 209];
- **медиана** (Me, если расставить тестовые баллы в один ряд, то срединное значение будет медианой);

- **мода** ( $M_o$ , наиболее повторяющееся значение тестовых баллов).

Чем лучше сделан тест и чем больше испытуемых, тем больше гистограмма похожа на кривую нормального распределения, известную в статистике. Чем больше показатели  $M$ ,  $M_e$ ,  $M_o$  отклоняются одна от другой, тем больше данные отклоняются от нормального распределения.

**Педагогический тест** (сокр. «тест») — репрезентативная педагогическая система параллельных заданий возрастающей трудности, специфической формы, позволяющая качественно оценить структуру и измерить уровень подготовленности испытуемых.

### ***9.1. Ориентиры при конструировании педагогического теста:***

Каждый тест имеет оптимальное время тестирования. Уменьшение или превышение которого снижает качественные показатели теста. Оптимальное время тестирования определяется эмпирически, по показателю дисперсии тестовых данных. Максимум значения последней укажет на оптимум времени.

Рекомендуемое время для выполнения одного тестового задания — не более 1.5-2 мин.

Рекомендуемое время и количество заданий для проведения тестирования:

- домашней работы по теме — 5-10 заданий (10-15 мин.)
- промежуточного по модулю — 15-25 (30-40 мин.)
- итогового — 30-40 (60 мин.)

Оптимальное время выполнения теста — 40-60 мин.

Оптимальное число заданий в тесте – примерно 30 заданий.

Полученный тест желательно обсудить с коллегами (1 – 3 чел.)

### ***9. 2. Проектирование одинаковых условий тестирования для всех студентов, что включает в себя:***

- общая инструкция;
- одинаковые условия для тестирования (место, обстановка, качество компьютерной техники и т.п.);
- время выполнения отдельных заданий или всего теста должно быть ограничено и быть общим для всех студентов;
- содержание заданий и уровень их трудности должен быть одинаков для всех студентов (параллельные варианты теста должны пройти предварительную эмпирическую и статистическую проверку, которая должна подтвердить их параллельность).

## **Приложения**

### **Приложение 1. Рекомендации по технологии проектирования дидактических тестов<sup>1</sup>**

#### **Теоретический этап**

##### **1. Цели тестирования**

---

<sup>1</sup> По материалам пособия [Дидактические тесты, 2004, 30].

2. Выбор подхода к созданию теста
3. Изучение учебного материала

#### **Практический этап 1**

4. Определение структуры теста (спецификация теста)
5. Разработка тестовых заданий
6. Экспертиза тестовых заданий
7. Корректировка тестовых заданий
8. Конструирование теста для апробации (с. 52)
  - упорядочение заданий для испытания теста
  - определение трудности заданий и их места в тесте
  - установление количества заданий в тесте
  - определение темпа выполнения теста
  - создание композиции тестовых заданий в дидактическом тесте
9. Разработка инструктивно-методического обеспечения теста
  - 1.1. Подготовка процедуры проведения теста, в том числе:
    - бланка заданий
    - бланка ответов
    - инструкции для студента
    - инструкции для преподавателя
    - ключей тестов
  - 1.2. Сценарий проведения тестирования
10. Экспертиза теста

#### **Экспериментальный этап**

11. Предварительное тестирование
12. Анализ и интерпретация результатов тестирования (определение качественных характеристик тестов)

#### **Практический этап 2**

13. Переработка теста на основе результатов предварительного тестирования
14. Составление окончательного варианта теста
15. Стандартизация теста (при необходимости)

## **Приложение 2. Расчет коэффициента корреляции баллов испытуемых по заданию с баллами по всему тесту (корреляция задания с критерием)<sup>1</sup>**

Из множества известных методов расчета приведем формулу наиболее часто применяемого метода, хотя и не совсем отвечающего строгим требованиям статистической теории<sup>2</sup>.

Для расчета  $r_{xy}$  формируется два так называемых вектор-столбца, один из которых - задание ( $X_j$ ), другой - критерий ( $Y$ ). Между значениями этих двух векторов и устанавливается мера связи, если таковая существует. Мера связи определяется посредством расчета коэффициента корреляции  $r_{jy}$  где символом  $r$  обозначается так называемый классический коэффициент корреляции Пирсона,  $j$  представляет номер коррелируемого задания, а символ  $Y$  - числовой вектор-столбец тестовых баллов испытуемых.

<sup>1</sup> Цитируется по [Аванесов, 2002, 167—170].

<sup>2</sup> различия со значениями теоретически более подходящих, т.н. бисериальных коэффициентов для практиков часто оказываются несущественными.

Формулы для расчета коэффициентов корреляции и примеры такого расчета даются ниже. Проверим, например меру связи ответов испытуемых по заданию №7 с суммой баллов испытуемых по всему тесту. Для этого строится вспомогательная таблица 3, в которой использованы соответствующие данные табл.2.

В колонке  $X_7$  приводятся значения баллов, полученных испытуемыми в седьмом задании. Сумма этих баллов равна 5.

Во второй колонке представлены тестовые баллы ( $Y_i$ ); в таблице представлено без индекса  $i$ , что позволяет не перегружать формулы;  $\sum Y_i = 65$ .

В третьей колонке даются произведения баллов каждого испытуемого по седьмому заданию ( $X_7$ ) и по сумме баллов ( $Y$ );  $\sum X_7 Y = 34$ . Это сумма попарных произведений  $X$  и  $Y$ .

В четвертой и пятой колонках - квадраты значений  $X_7$  и  $Y$ ; Соответственно,  $\sum X_7^2 = 5$  и  $\sum Y^2 = 387$ .

При расчете используются, последовательно, четыре формулы:

1) Вначале находится сумма квадратов отклонений баллов испытуемых от среднего арифметического балла в интересующем задании ( $SS_7$ ). Это делается по уже знакомой формуле:

$$SS_7 = \sum X_7^2 - \frac{(\sum X_7)^2}{N} = 5 - \frac{5^2}{13} = 3,077$$

Расчет коэффициента корреляции.

Испытуемые	$X_7$	$Y$	$X_7 Y$	$X^2$	$Y^2$
1.	1	9	9	1	81
2.	1	8	8	1	64
3.	1	7	7	1	49
4.	0	6	0	0	36
5.	0	6	0	0	36
6.	1	5	5	1	25
7.	1	5	5	1	25
8.	0	5	0	0	25
9.	0	4	0	0	16
10.	0	4	0	0	16
11.	0	3	0	0	9
12.	0	2	0	0	4
13.	0	1	0	0	1
$\Sigma$ :	5	65	34	5	387

2) Затем находится сумма квадратов отклонений тестовых баллов испытуемых от среднего арифметического балла по всему тесту ( $SS_y$ ). Это делается по формуле:

$$SS_y = \sum Y^2 - \frac{(\sum Y)^2}{N} = 387 - \frac{4225}{13} = 62.$$

3) Находится так называемая скорректированная на средние значения сумма попарных произведений  $X$  и  $Y$ , по формуле:

$$SP_{xy} = \sum XY - \frac{\sum X \sum Y}{N} = 34 - \frac{5 \times 65}{13} = 9$$

В последней формуле  $\sum XY$  представляет собой сумму произведений баллов каждого испытуемого по седьмому заданию и по  $Y_i$ , тестовому баллу испытуемых. Вторая часть формулы  $\frac{\sum X \sum Y}{N}$  представляет собой коррекцию на средние значения  $X$  и  $Y$ .

4) Рассчитывается коэффициент корреляции по формуле:

$$r_{xy} = \frac{SP_{xy}}{\sqrt{SS_x \times SS_y}} \quad (11.8)$$

Подставляя в эту формулу результаты проведенных расчетов, получаем

$$\frac{9}{\sqrt{3,077 \times 62}} = 0,652$$

$r_{xy} =$

### Приложение 3. Словарь основных понятий, используемых в Руководстве<sup>1</sup>

**Тестовая форма (педагогического) задания** — характеризуется следующими признаками: точностью и адекватностью выражения определенного типа содержания, понятностью для всех испытуемых, невозможностью появления ошибочных ответов по содержательным и формальным признакам.

**Нетестовая форма (педагогического) задания** — задачи, вопросы, примеры, упражнения, кроссворды и пр. В тест они не включаются по причинам нетехнологичности, громоздкости, затруднений в восприятии формы и содержания, повышенной трудоемкости и повышенным временным затратам. Это приводит к снижению эффективности и качества измерений, а часто и к невозможности измерения вообще.

**Вопрос** — нетестовая форма педагогического задания (наряду с задачами, учебными проблемами, упражнениями и т.д.). Смысл тестового утверждения, излагаемого в простой структуре предложения в повествовательной форме улавливается всегда лучше, чем смысл вопроса. Для верного понимания смысла вопроса очень важен контекст. В тестовых утверждениях нет ни одного лишнего слова и даже знака, в то время как вопрос требует ряда дополнительных слов и знаков для выражения требуемого смысла, значения и интонации. Вопросы и ответы на них иногда бывают столь неопределенными и многословными, что для выявления истинности требуются большие затраты интеллектуальной энергии. Учебные вопросы порождают ответы, полные и неполные, правильные и неправильные, разные по форме, содержанию и по структуре, вследствие чего оценка таких ответов требует обязательного участия преподавателя и сопровождается некоторой долей субъективизма. В то время как технологичная методика тестирования предполагает четкую и быструю дифференцируемость правильного ответа от неправильного. В этом смысле традиционные вопросы и ответы не технологичны; их не рекомендуется включать в тест. То же относится и к задачам, имеющим тяжеловесные формулировки. Логическое преимущество задания в тестовой форме заключается в возможности естественного превращения утверждения, после ответа студента, в форму истинного или ложного высказывания.

**Задание в тестовой форме** — вид педагогического задания, которое характеризуется стандартизированной, т. е. **тестовой**, формой фиксации — во-первых, указаний выполнить то или иное действие студенту и, во-вторых, результатов проделанной им работы (в узком смысле — ответов).

Задание в тестовой форме — это первое основное понятие педагогической теории измерений. Задание в тестовой форме (в отличие от нетестовой) отвечает требованиям (более подробно см. в тексте Руководства):

- цель
- краткость
- технологичность
- правильность формы
- логическая форма высказывания
- одинаковость правил оценки ответов
- наличие определенного места для ответов
- правильность расположения элементов задания
- одинаковость инструкции для всех испытуемых
- адекватность инструкции форме и содержанию задания.

<sup>1</sup> При составлении данного словаря мы ориентировались на подход и работы В.С.Аванесова.

Наиболее распространенной является такой тип заданий (**закрытые**) в тестовой форме, когда варианты решения задания (результаты действия) предлагаются на выбор студенту. Среди них есть правильные(ый) и неправильные. Студент должен определиться в отношении предложенных ему вариантов. Иным типом являются **открытые задания**, в которых готовые решения (на выбор) не даются. Студент сам должен сформулировать решение (ответ) и зафиксировать его в специально отведенном месте.

**Дистрактор** — неправильный, но правдоподобный ответ (от англ. to distract — отвлекать).

**Типы заданий в тестовой форме («формы тестовых заданий» по В.С.Аванесову).**

В Руководстве используется классификация В.С.Аванесова, который выделяет всего 4 известных на сегодня типа заданий: **с выбором одного или нескольких правильных ответов; на установление правильной последовательности; на установление соответствия; открытой формы**. В Руководстве мы используем термин «тип» для классификации заданий в тестовой форме, относя понятие «тестовой формы» к стандартизированному краткому виду предъявления небольших заданий. Кроме того в лексике В.С.Аванесова невозможно сказать — форма задания в тестовой форме, а выражение «форма тестовых заданий», которое он использует, затрудняет понимание, потому что чаще всего относится к «заданиям в тестовой форме», а не к «тестовым заданиям», которые он строго различает.

Каждый тип проверяет в той или иной степени разные виды знаний, поэтому их выбор зависит от учебного предмета, цели применения теста, способа обработки данных и, в целом, является творческим выбором самого преподавателя.

**Набор заданий в тестовой форме** (НЗТФ, данного понятия в словаре В.С.Аванесова нет) — простая исходная совокупность отдельных заданий в тестовой форме, не прошедшая экспериментальной проверки и служащая материалом промежуточной стадии при разработке теста. Для каждого учебного модуля необходимо разработать порядка 100-150 «заданий в тестовой форме» для того, чтобы на следующем этапе можно было бы отобрать из них «тестовые задания» (около половины). Полученный НЗТФ должен быть «размножен» до нескольких вариантов (В.С.Аванесов рекомендует иметь до 15 вариантов). Это поможет сконструировать параллельные варианты теста.

Шансы попасть в тест имеют только те задания в тестовой форме, у которых правильная форма, точно сформулированное и корректно отобранное содержание, а также есть математико-статистические характеристики, необходимые для создания теста как педагогической системы.

**Система заданий в тестовой форме** (СЗТФ) — надо отличать от промежуточной стадии при разработке теста — «набора заданий в тестовой форме» (НЗТФ). *СЗТФ образуют такие задания, которые имеют все признаки, присущие заданиям в тестовой форме, но не обладают свойствами тестовых заданий, а поэтому они в принципе тестом не являются.* СЗТФ применяют для педагогического оценивания, не претендующего на соответствие строгим требованиям научного измерения, что свойственно тестам. При применении СЗТФ получаемые студентами баллы сводятся к обычным оценкам (в результате субъективного оценивания баллов), которые не могут интерпретироваться как результат педагогического измерения.

К СЗТФ относятся **цепные** (правильный ответ на последующее задание зависит от ответа на предыдущее задание), **тематические** (совокупность заданий любой формы, созданная для контроля по одной изученной теме), **текстовые** (совокупность заданий открытой формы, созданных для автоматизированного контроля знаний учебного текста), **ситуационные** (педагогически переработанный фрагмент профессиональной деятельности специалиста; разрабатываются для проверки знаний и умений испытуемых действовать в практических, ситуациях, включая экстремальные; дается текстовой описание ситуации и к ней прикладывается система заданий в тестовой форме). **В СЗТФ реализуется принцип — вероятность правильного ответа на последующее задание может зависеть от**

*вероятности правильного ответа на предыдущие задания. Это нарушает принцип статистической независимости тестовых заданий (аксиома локальной независимости), который положен в основу создания теста как формальной системы. Поэтому задания в таких системах (СЗТФ) теста не образуют.*

**Тестовое задание** — это *составная единица* теста, отвечающая всем требованиям к заданиям в тестовой форме (технологичности, чистоты формы и предметного содержания) и, кроме того, *статистическим требованиям*:

- 1) известной (меры) трудности;
- 2) дифференцирующей способности (достаточной вариации тестовых баллов);
- 3) положительной корреляции баллов испытуемых по заданию с баллами по всему тесту (сокращ. — корреляция задания с критерием)
- 4) а также другим математико-статистическим требованиям, которые в данной инструкции опускаются («подходящий геометрический образ задания» и т.д.).

Из перечисленных требований вытекает обязательность эмпирической проверки заданий на выборочной совокупности испытуемых и применения статистических методов обработки данных.

**Тестовый балл испытуемого** — суммирование всех баллов, полученных испытуемым, дает число правильных ответов. Это число ассоциируется с уровнем его знаний.

**Педагогический тест** (сокр. «тест») — репрезентативная педагогическая система параллельных заданий возрастающей трудности, специфической формы, позволяющая качественно оценить структуру и измерить уровень подготовленности испытуемых.

Тест надо рассматривать как единство: 1) метода педагогического измерения; 2) результатов, полученных определенным методом (матрица данных); 3) интерпретированных результатов, полученных определенным методом. Тест — это творческое соединение метода и правильно интерпретированных результатов его применения.

Качественный тест образует шкалу, которая может быть определена как средство для измерения свойств объектов.

Все три понятия — задание в тестовой форме, тестовое задание и тест — соответствуют *трем основным этапам разработки педагогических тестов*:

1. Разработка заданий в тестовой форме.
2. Создание тестовых заданий (статистическая проверка заданий, с целью отбора их в качестве тестовых заданий).
3. Моделирование, конструирование теста из тестовых заданий.

Тест состоит не из заданий в тестовой форме, не из вопросов и задач, а только из тестовых заданий.

**Параллельные варианты теста** — имеют сходное предметное содержание, равные средние арифметические, равные дисперсии и равные интеркорреляции. Параллельный вариант теста, имея внешне отличающееся содержание, должен, концептуально, измерять то же самое, что и исходный тест, с той же точностью.

**Предметно-педагогическое тестирование** (интерпретация тестовых результатов) — направлено на определение того, какие элементы содержания учебной дисциплины усвоены данным испытуемым (содержательная интерпретация)

**Критериально-ориентированное тестирование** (интерпретация тестовых результатов) — направлено на определение того, соответствуют ли знания испытуемого зачетному уровню трудности (критерию, критериальная интерпретация).

**Нормативно-ориентированное тестирование** (интерпретация тестовых результатов) — направлено на определение сравнительного места, рейтинга испытуемых относительно друг друга в определенной содержательной области (при этом используется статистический язык нормального распределения результатов тестирования и понятие процентильных норм, показывающих — сколько процентов испытуемых имеют тестовый результат хуже, чем у любого взятого для анализа испытуемого, с его тестовым баллом).



**Эффективность теста** — если он лучше, чем другие измеряет знания студентов интересующего уровня подготовленности, с меньшим числом заданий, качественнее, быстрее, дешевле, и все это — по возможности, в комплексе. Эффективный тест состоит из эффективных заданий, которые эксперты признают в качестве ключевых элементов изучаемой дисциплины. В принципе возможна ситуация, когда из большего числа заданий можно сделать такую выборку, которая не уступает первой. Тест с меньшим числом заданий — более эффективен, также как и тест, точно соответствующий по трудности заданий уровню подготовленности испытуемых. Поэтому тест не может быть эффективным вообще, на всем диапазоне подготовленности студентов.

**Качество педагогического теста** — традиционно сводится к определению меры надежности и валидности полученных результатов.

**Валидность тестовых результатов** — одно из двух основных критериев качества тестовых результатов; означает меру пригодности результатов для определенной цели тестирования; зависит от качества заданий, их числа, от степени полноты и глубины охвата содержания учебного модуля (дисциплины) в заданиях теста; зависит также от баланса и распределения заданий по трудности, от метода отбора заданий в тест из общего банка заданий, от интерпретации тестовых результатов, от организации сбора данных, от подбора выборочной совокупности испытуемых.

**Надежность тестовых результатов** — одно из двух основных критериев качества тестовых результатов; в последние годы стали чаще писать о надежности измерения, имея в виду тест как результат тестирования в конкретных условиях, а не общую характеристику теста как метода; в практическом смысле надежность обычно понимается как мера одинаковости, повторяемости или связанности двух измерений одного и того же качества, одним и тем же тестом, или его параллельными вариантами, часто выражается мерой корреляции между двумя **параллельными вариантами теста**. Надежность тестовых результатов — это устойчивость (стабильность) индивидуальных тестовых баллов. В случае опоры на классическую теорию тестов главный метод — использование параллельных вариантов тестов, с последующим коррелированием результатов. Получаемое при этом значение интерпретируется в смысле воспроизводимости (близости или параллельности) тестовых результатов. В этой функции используется и повторное тестирование испытуемых, в одинаковых условиях, одним и тем же тестом, с последующим коррелированием результатов. Малое число испытуемых и малое число заданий дают ненадежные результаты тестирования. Существуют многочисленные методы определения надежности тестовых результатов.

## Литература

### Список цитируемой литературы

*Аванесов, 2002*: Аванесов В.С. Композиция тестовых заданий. Учебная книга. 3-е изд., доп. — М.: Центр тестирования, 2002. — 240 с.

*Аванесов, 2005*: Аванесов В.С. «Формы тестовых заданий». Учебное пособие для учителей школ, лицеев, преподавателей вузов и колледжей. 2-е изд. перераб. и расширен. — М.: Центр тестирования, 2005. — 156 с.

*Дидактические тесты, 2004*: Дидактические тесты: технология проектирования: методическое пособие для разработчиков тестов /Е.В. Кравец, А.М. Радьков, Т.В.Столярова, Б.Д. Чеботаревский; Под общ. ред. А.М.Радькова. — Мн.: РИВШ, 2004. — 87 с.

*Жук, 2005*: Жук О.Л. и др. Педагогические основы самостоятельной работы студентов: Пособие для преподавателей и студентов / Под общ. Ред. О.Л. Жук. — Мн.: РИВШ, 2005. — 112 с.

*Трофимова, 2005*: Трофимова З.П. Основы методологии и методики построения педагогических тестов: учеб. - метод. пособие / Под ред. А.В.Макарова. — Мн: РИВШ, 2005. — 60 с.

*Равен, 1999*: Равен Джон Педагогическое тестирование: Проблемы, заблуждения, перспективы / Пер. с англ. — М.: «Когито-Центр», 1999. — 144 с.

*Равен, 2002*: Равен Дж. Компетентность в современном обществе: выявление, развитие и реализация / Пер. с англ. — М.: «Когито-Центр», 2002. — 396 с.

### Список рекомендуемой литературы

Майоров А.Н. Теория и практика создания тестов для системы образования. (Как выбирать, создавать и использовать тесты для целей образования). — М.: Интеллект-центр, 2001. — 296 с.

Михайлычев Е.А. Дидактическая тестология. — М.: Народное образование, 2001. — 432 с. (Серия «Профессиональная библиотека учителя»)

Чельшкова М. Б. Теория и практика конструирования педагогических тестов: Учебное пособие. — М.: Логос, 2002. — 432 с.

Тесты проверки знаний: этапы разработки / сост. Н.П.Радчикова. — Минск : РИВШ, 2007. — 30 с.