

Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ



Tempus

Е.Г. Григорьева, Н.Л. Гагулина, А.В. Черемных

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОЦЕНКЕ ПРОФЕССИОНАЛЬНЫХ КВАЛИФИКАЦИЙ



РГГМУ

Санкт-Петербург

2012

УДК 378.146

ББК 74.58в6

Г - 83

Григорьева Е.Г.

Методические указания по оценке профессиональных квалификаций: Учебно-методическое пособие / Е.Г. Григорьева, Н.Л. Гагулина, А.В. Черемных. – СПб., 2012. – 64 с.

В пособии рассматриваются вопросы применения тестов для оценки уровня профессиональных квалификаций, в том числе и для проведения текущего, рубежного и итогового контроля качества полученных знаний и приобретенных компетенций у студентов и выпускников образовательных программ. Содержатся практические рекомендации по разработке тестовых заданий различной формы, приводится большое количество примеров формирования тестовых заданий по направлению «прикладная гидрометеорология». Адресовано преподавателям вуза, разрабатывающим тесты, методическим комиссиям, проводящим экспертизу тестов, специалистам, организующим проведение тестирования, а также для потребностей УМО в области гидрометеорологии. Будет полезна всем, кто интересуется вопросами тестирования и оценкой уровня профессиональных квалификаций.

Проект "Разработка структуры квалификационных требований в метеорологии" (TEMPUS JPHES-159352 QUALIMET) финансируется при поддержке Европейской Комиссии.

Содержание данной публикации отражает мнение авторов и Европейская Комиссия не несет ответственности за использование содержащейся в ней информации.

Табл. 6. Ил. 4. Библиогр.: 11 назв.

© Е.Г. Григорьева, Н.Л. Гагулина, А.В. Черемных, 2012

© Российский государственный

гидрометеорологический университет, 2012

СОДЕРЖАНИЕ

СЛОВАРЬ ТЕРМИНОВ	5
ВВЕДЕНИЕ	9
1 ОБЩИЕ ТРЕБОВАНИЯ К ТЕСТАМ	12
1.1 Классификация тестов	12
1.2 Основные принципы разработки тестов	15
2 ЭТАПЫ РАЗРАБОТКИ ТЕСТА ДЛЯ АТТЕСТАЦИИ СТУДЕНТОВ	16
2.1 Перечень этапов и их очередность	16
2.2 Разработка плана теста и спецификации	17
3 ФОРМЫ ТЕСТОВЫХ ЗАДАНИЙ	20
3.1 Общие требования к ТЗ	20
3.2 Правила оценки тестовых заданий и теста	21
3.3 Классификация форм ТЗ.....	24
3.4 Задания закрытой формы с выбором ответа.....	25
3.4.1 Общие требования к тестовым заданиям закрытой формы с выбором ответа.....	25
3.4.2 Задания альтернативного выбора.....	26
3.4.3 Задания с выбором одного варианта из набора.....	27
3.4.4 Задания с выбором нескольких вариантов из набора....	28
3.5 Задания на установление соответствия.....	29
3.6 Задания на установление правильной последовательности.	31
3.7 Задания открытого типа.....	32
3.8 Сравнительная характеристика форм ТЗ	34
3.9 Составление заданий с разным принципом композиции	36
3.10 Кейс-задания.....	41
3.11 Норма трудности тестовых заданий.....	44
4 АНАЛИЗ КАЧЕСТВА СОСТАВЛЕНИЯ ТЕСТОВ И ТЕСТОВЫХ ЗАДАНИЙ	46

4.1 Общие операции.....	46
4.2 Расчет качественных характеристик теста	47
4.3 Определение качественных характеристик тестовых заданий	50
ЗАКЛЮЧЕНИЕ	53
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	54
ПРИЛОЖЕНИЕ А. ПРИМЕРЫ СПЕЦИФИКАЦИИ ТЕСТА ПО ДИСЦИПЛИНЕ «ОБЩАЯ ОКЕАНОЛОГИЯ»	55
ПРИЛОЖЕНИЕ Б. ПРОЕКТ ЭКСПЕРТНОГО ЗАКЛЮЧЕНИЯ ЭКСПЕРТИЗЫ ТЕСТОВЫХ ЗАДАНИЙ	58
ПРИЛОЖЕНИЕ В. ПРИМЕР СЕРИИ АЛЬТЕРНАТИВНЫХ ВОПРОСОВ	60
ПРИЛОЖЕНИЕ Г. ПРИМЕРЫ ОШИБОК В ЗАДАНИЯХ	61

СЛОВАРЬ ТЕРМИНОВ

Апробация теста – пробное педагогическое тестирование, предшествующее основному и предпринимаемое с целью устранения ошибок и неточностей в заданиях, а также приблизительной оценке трудности теста и его отдельных частей (заданий).

База тестовых заданий – совокупность тематически связанных заданий, из которых формируется тест.

Банк тестовых заданий – совокупность тестовых заданий или тестов с устойчивыми характеристиками, представленными в компьютерной форме.

Балл – условная единица для оценки по определенной шкале результатов выполнения теста или его задания.

Валидность теста – комплексная характеристика теста, отражающая его способность измерять именно то, для чего он предназначен. Характеризует возможности генеральной совокупности заданий в тестируемой области знаний оценить объект измерений теста.

Градуированная шкала – шкала, отметки которой должны с требуемой точностью соответствовать значениям измеряемой величины.

Дидактическая единица (ДЕ) – целостный раздел (модуль, тема учебной дисциплины).

Дистрактор (отвлекающий ответ) – вариант ответа на тестовое задание закрытого типа, похожий на правильный, но не являющийся таковым.

Длина теста – количество тестовых заданий теста.

Задание закрытого типа – тестовое задание, содержание которого сопровождается несколькими номерованными вариантами ответа; испытуемому предлагается выбрать номер или номера правильных ответов.

Задание альтернативного типа – тестовое задание закрытого типа; испытуемому предлагается выбрать правильный ответ из двух предложенных (да/нет).

Задание с выбором одного варианта из нескольких – тестовое задание закрытого типа; испытуемому предлагается выбрать правильный один ответ из набора предложенных.

Задание с множественным выбором – тестовое задание закрытого типа; испытуемому предлагается выбрать несколько правильных ответов из предложенных вариантов.

Задание открытого типа – тестовое задание без указания возможных вариантов ответа; испытуемому предлагается самостоятельно указать правильный ответ.

Задание с коротким регламентированным ответом (на дополнение) – тестовое задание открытого типа; испытуемому предлагается дополнить незаконченное высказывание до получения истинного или ложного утверждения.

Задание со свободно конструируемым ответом – тестовое задание открытого типа; испытуемому предлагается самостоятельно сформулировать ответ; никакие ограничения на него в задании не накладываются.

Задание на соответствие (тестовое задание перекрестного выбора) – тестовое задание, при выполнении которого необходимо установить соответствие между элементами двух множеств (двух списков).

Задание на установление правильной последовательности – тестовое задание, выполнение которого состоит в установлении правильной последовательности операций, действий, событий.

Инструкция тестового задания – словесные указания испытуемому, связанные с выполнением тестового задания (выбором правильного ответа из нескольких вариантов; решением математической задачи и т.п.). Указывается способ записи правильного ответа (что, каким образом и где надо отметить, вписать и т.д.). Инструкция может быть единой для нескольких заданий теста, если эти задания однотипны по действиям.

Кейс-задание – задание, состоящее из описания конкретно-практической ситуации и совокупности вопросов (задач) к ней.

Компетентностно-ориентированные задания – задания, решения которых требует комплексной деятельности по решению проблемы (поиск информации, обмен информацией, структурирование и применение информации, работа в группе, вербальные и невербальные способы коммуникации, подготовка презентаций и защита результатов деятельности).

Компетентностный тест – тест, направленный на измерение уровня формирования компетенций выпускников и предна-

значенный для итоговой государственной аттестации выпускников. Состоит из компетентно-ориентированных заданий.

Компетентность – совокупность компетенций, наличие знаний и опыта, необходимых для эффективной деятельности в заданной предметной области.

Компетенция – способность применять знания, умения и личностные качества для успешной деятельности в определенной области.

Мультимедиа – совокупность компьютерных технологий, использующих графику, текст, видео, фотографию, анимацию, звуковые эффекты. Технологию мультимедиа обеспечивают специальные аппаратные и программные средства.

Надежность теста – показатель точности и устойчивости результатов измерения с помощью теста при его многократном применении. Характеризует степень адекватности отражения тестом соответствующей генеральной совокупности заданий.

План теста – таблица, в которой каждое тестовое задание соотносится с определенным элементом содержания учебного предмета, конкретным видом знаний или умений, позволяющим объективно судить об уровне подготовленности испытуемых. Указывается также планируемое время выполнения каждого тестового задания и всего теста в целом.

Спецификация теста – документ, в котором содержится информация о целях, задачах, плане и структуре теста, а также указаны основные требования к правилам проведения тестирования, обработки результатов тестирования и их интерпретации.

Структура теста – совокупность сведений о числе и типе тестовых заданий по каждому элементу содержания учебного предмета и по каждому виду знаний или умений, позволяющих объективно судить об уровне подготовленности испытуемых. Указывается также предполагаемый уровень трудности каждого тестового задания и, по возможности, его коэффициент дискриминации.

Тест педагогический – система заданий специфической формы, позволяющая качественно оценить структуру и измерить уровень достижений обучаемых. Предполагает формализованные ответы испытуемого при выполнении формализованных заданий.

Тест нормативно-ориентированный – система стандартизированных тестовых заданий, ранжированных в рамках определен-

ной стратегии предъявления и позволяющих эффективно дифференцировать испытуемых по уровню и качеству подготовки.

Тест критериально-ориентированный – система стандартизированных однородных по содержанию тестовых заданий, различной трудности, предназначенных для итогового контроля и позволяющих оценить процент освоенных знаний и умений по отношению к полному объему.

Тестирование аттестационное (итоговое) – педагогическое тестирование по окончании обучения, с целью сопоставления уровня подготовленности студентов с образовательными стандартами ФГОС ВПО и аттестации испытуемых.

Тестирование тематическое – педагогическое тестирование с целью измерения уровня подготовленности испытуемых в определенной области знаний, соответствующей законченной теме или разделу учебной дисциплины.

Тестовое задание (ТЗ) – задание специфической формы, элемент теста, минимальная законченная составляющая единица теста.

Фонд оценочных средств – комплект методических и контрольных измерительных материалов, предназначенных для оценивания достижений и компетенций студентов на разных стадиях обучения, материалы аттестационных испытаний выпускников на соответствие (или несоответствие) уровня их подготовки требованиям соответствующего ФГОС ВПО по завершению освоения ООП по определенному направлению или специальности.

Форма (тестового) задания – общая схема составления тестового задания, его дизайн и структура.

Шкала – упорядоченное множество действительных чисел (индексов), соответствующих ряду возможных значений измеряемой величины.

ВВЕДЕНИЕ

Необходимость оценки профессиональных квалификаций имеет место на всех стадиях образовательного процесса: от входной аттестации, определяющей базовый уровень подготовки студентов к освоению содержания основной образовательной программы высшего профессионального образования, через все виды промежуточной аттестации и до итоговой аттестации, которая ставит точку на освоении программы в пределах высшего учебного заведения и открывает двери перед выпускником, готовым приступить к профессиональной деятельности в рамках присвоенной квалификации.

Введение нового поколения ФГОС в практику высшего профессионального образования повышает актуальность оценки профессиональных квалификаций. В стандарте ФГОС по прикладной гидрометеорологии определено [1]: *«Для аттестации обучающихся на соответствие их персональных достижений поэтапным требованиям соответствующей ООП (текущий контроль успеваемости и промежуточная аттестация) создаются фонды оценочных средств, включающие типовые задания, контрольные работы, тесты и методы контроля, позволяющие оценить знания, умения и уровень приобретенных компетенций. Фонды оценочных средств разрабатываются и утверждаются вузом».*

Создаваемые вузом фонды оценочных средств (ФОС) должны проходить многокомпонентную экспертизу:

- вузовскую;
- УМО вузов (НМС) по направлениям и специальностям;
- представителей работодателей и профессиональных сообществ;
- Рособнадзора.

Экспертиза фондов оценочных средств должна осуществляется на предмет выявления:

- соответствия ФОС структуре требований ФГОС;
- соответствия ФОС образовательной программе по направлению подготовки;
- соответствия результатов обучения задачам профессиональной деятельности.

Таким образом, ФОС должны полностью отражать требования ФГОС ВПО и обеспечивать решение оценки соответствия универсальных и профессиональных компетенций (квалификационных требований) выпускника профессиональным стандартам.

Вузом самостоятельно проводится проектирование оценочных средств для проведения текущего, промежуточного, рубежного контроля и итоговой аттестации обучаемых в соответствии с учебным планом вуза по направлению подготовки и учетом особенностей принятых образовательных программ.

Модель ФОС может состоять из традиционных и инновационных компонентов, которые выполняют функцию комплексной проверки и оценки самых разнообразных компетенций:

- собеседование;
- коллоквиум;
- зачет;
- экзамен (по дисциплине, модулю, итоговый государственный экзамен);
- тест;
- контрольная работа;
- эссе и иные творческие работы;
- реферат;
- отчет (по практикам, научно-исследовательской работе студентов и т.п.);
- курсовая работа;
- выпускная квалификационная работа;
- кейсы;
- портфолио.

По мнению авторов [2-11] тесты являются эффективным средством контроля результатов обучения на уровне теоретических знаний, понимания и умения применять знания на практике. Позволяют не только получить объективную информацию о качестве знаний и умений обучаемых, но и определить наиболее слабо усвоенные разделы, темы, отдельные вопросы и своевременно скорректировать процесс обучения. В связи с необходимостью объективной оценки компетенций выпускника применение тестов (как в случае тестовой формы экзамена, так и для других оценочных испытаний) должно сопровождаться процедурой оценки качества применяемых тестов. Тесты обязаны соответствовать определен-

ным критериям, и работа по их формированию должна осуществляться только квалифицированными преподавателями и специалистами в данной области. Возможно, вузу придется создавать специальное подразделение в своей структуре, которое займется процедурой тестирования, или приглашать профессиональных специалистов со стороны.

Для развития тестовых технологий образовательному учреждению важно провести работу по формированию *единых подходов* к проектированию тестов и созданию *фонда тестовых заданий* в рамках всех направлений подготовки. На данном этапе приоритетной задачей в этом направлении являются:

- обучение ППС университета по вопросам методики и технологии компьютерного тестирования;
- техническое оснащение тестового процесса;
- разработка большого числа заданий в тестовой форме для различных целей применения, особенно для использования в самостоятельной работе студентов;
- контроль качества заданий, которые планируется использовать для создания тестов.

Методическое пособие представляет сжатое, но достаточно полное описание теории и практики тестового контроля на основе современных представлений средств оценивания результатов обучения [3-10]. Является переработанным изданием [11]. В публикации используются термины и определения, представленные в словаре терминов. Математические выкладки статистической обработки тестов, рассмотренные в разделе 4, приводятся без выводов, т.к. представлены в использованных источниках [3, 5, 6, 8, 9].

Авторы подтверждают, что в отношении оценки профессиональных квалификаций, с точки зрения тестового контроля, в этом документе сделаны только первые шаги.

1 ОБЩИЕ ТРЕБОВАНИЯ К ТЕСТАМ

Для построения теста, как объекта проверки профессиональных квалификаций, необходимо определить цели и выбрать, соответственно им, подходы и область применения планируемого тестового контроля.

1.1 Классификация тестов

В литературных источниках рассмотрены теоретические подходы и приведены практические рекомендации к составлению тестов и тестовых заданий [3-10]. Авторами приводятся классификации педагогических тестов, в основе которых лежат различные подходы к их формированию и представлению. Обобщенная классификация тестов представлена на рисунке 1.

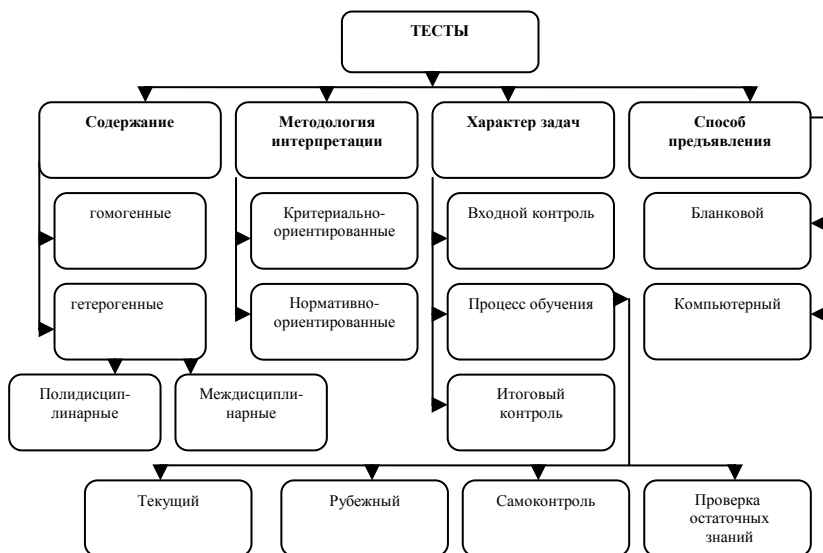


Рисунок 1 – Классификация тестов

По содержанию тесты делятся на гомогенные и гетерогенные. *Гомогенные* тесты измеряют уровень подготовленности по одной

учебной дисциплине, а *гетерогенные* тесты состояются из системы заданий по нескольким учебным дисциплинам. Гетерогенные бывают *полидисциплинарными*, которые состоят из однородных субтестов по отдельным дисциплинам и *междисциплинарные*, где требуется применение обобщенных межпредметных интегральных знаний и умений.

К наиболее важным целевым характеристикам тестов относится *методология интерпретации*: критериально-ориентированный или нормативно-ориентированный тест. *Критериально-ориентированные* тесты позволяют измерить уровень индивидуальных учебных достижений относительно полного объема знаний, навыков и умений, которые должны быть усвоены обучаемыми по конкретной дисциплине. Результаты контроля оцениваются по отношению к содержательной области или требованиям, установленным в ФГОС. Недостатки критериально - ориентированного подхода связаны с необходимостью полного охвата дисциплины, в результате чего длина теста может быть очень значительна. Тесты такого типа применяются в текущем, рубежном и итоговом контроле. *Нормативно-ориентированные* тесты позволяют сопоставлять учебные достижения (уровни профессиональных знаний и умений) отдельных испытуемых друг с другом. Примером нормативно-ориентированного теста является олимпиадный тест.

По характеру задач тесты делятся на группы:

а) *входные тесты* – предназначены для оценки определения уровня знаний, приобретённых на предшествующем этапе обучения.

б) тесты *процесса обучения* со следующими типами контроля:

– *оперативный* или *текущий* используется в повседневном учебном процессе для проверки отдельного раздела, темы;

– *рубежный* (промежуточной аттестации) проводится в конце учебного периода (семестра, курса) по всей дисциплине (или ее части);

– *самоконтроль* применяется в качестве обучающей технологии для самостоятельной отработки студентами темы, типа задач и т.д.;

– проверка *остаточных знаний* проводится при прохождении вузом процедуры самообследования для государственной аккредитации, контролирующее задания для их оценки формируются согласно рабочей программе дисциплины и требованиям ФГОС.

в) тесты *итоговой аттестации* (междисциплинарный экзамен, предварительный этап госэкзаменов) применяются как результирующее тестирование в конце обучения: проверка уровня готовности специалиста, его соответствие требованиям ФГОС.

По способу предъявления тесты можно разделить на *компьютерные* и *бланковые*. Возможности компьютерных тестов шире, т.к. они ориентированы на применение информационных технологий. Современные программы для сопровождения тестового контроля обладают высокой надёжностью, позволяют экономить время за счет скорости проверки качества усвоения знаний, обеспечивают высокую вариантность тестов и имеют возможность удаленно работать с базой тестовых заданий. Кроме этого, они могут поддерживать мультимедийный режим. *Компьютерные* тесты делятся на:

– линейные – задаются все задания выборки, затем подсчитывается оценка;

– адаптивные – строятся на модели обучаемого: задания предъявляются в порядке, зависящем от ответа испытуемого на предыдущее задание, что позволяет с помощью разных стратегий предъявления заданий значительно сократить время тестирования.

Бланковые тесты традиционны, удобство заключается в проведении контроля непосредственно на занятии в обычной аудитории.

По характеру *измеряемых достижений* выделяют тесты для проверки навыков, учебных умений, практических умений, а также *компетентностные* тесты, направленные на измерение уровня формирования компетенций, и предназначенные для итоговой государственной аттестации выпускников.

Тесты можно разделить на *стандартизированные* тесты, в которых заданы нормы выполнения, и *нестандартизированные*, которые используются в повседневном учебном процессе.

1.2 Основные принципы разработки тестов

Разработка тестовых материалов для любой формы контроля (бланкового или компьютерного) должна удовлетворять определенным принципам, выработанным требованиями, правилам оформления и требованиям по оценке тестовых заданий. При разработке тестовых заданий необходимо соблюдать ряд принципов:

- *Соответствие* содержания теста *целям тестирования*.
- *Соответствие* содержания тестовых заданий (ТЗ) *программе учебной дисциплины и требованиям ФГОС*.
- *Полнота и значимость* материала, для которого разработаны ТЗ. В тест включаются наиболее значимые элементы содержания, относящиеся к опорным темам курса.
- *Системность* материала, что предполагает подбор упорядоченных содержательных элементов, связанных между собой определенной иерархией и общей структурой знаний.
- *Научность* материала тестирования. Содержание теста должно соответствовать уровню современного состояния науки. В тест включается только то содержание учебной дисциплины, которое является объективно истинным и поддается аргументации.
- *Достоверность* тестирования, т.е. адекватность результатов проверки изученному материалу.
- *Вариативность* содержания теста, т.е. возможность замены элементов задания другими похожими элементами, что обеспечит многовариантность контроля.
- *Объективность* оценки результатов тестирования, что исключает проявление субъективного подхода к результатам обучения.
- *Разнообразие форм* тестовых заданий, позволяющих всесторонне и адекватно оценить уровень подготовленности обучаемого и выбор категории трудности.

Для компьютерного контроля добавляются принципы:

- *Доступность* тестового контроля, что является необходимым условием для дистанционного и открытого обучения.
- *Технологичность* тестирования - применение автоматизированного сбора, обработки хранения тестов и результатов тестирования, что обеспечивает условия создания единой информационно-образовательной среды.

2 ЭТАПЫ РАЗРАБОТКИ ТЕСТА ДЛЯ АТТЕСТАЦИИ СТУДЕНТОВ

2.1 Перечень этапов и их очередность

Процесс создания аттестационных тестов по дисциплине (к которым относятся и тесты промежуточной аттестации), его научного обоснования, переработки и улучшения можно разбить на несколько этапов. В частности, необходимо:

1. Определить подход к созданию аттестационного теста (нормативно или критериально-ориентированный).
2. Провести анализ содержания учебной дисциплины, выделить разделы содержания дисциплины и темы заданий в каждой ДЕ, что позволит определить репрезентативное отображение требований ФГОС в тесте.
3. Определить структуру теста и стратегию расположения заданий.
4. Разработать спецификацию теста, осуществить априорный выбор длины теста, форм его заданий и времени его выполнения.
5. Создать задания в каждой из выбранных тестовых форм.
6. Отобрать задания в тест и ранжировать их согласно выбранной стратегии предъявления.
7. Провести экспертный анализ качества содержания заданий, теста и формы тестовых заданий.
8. По результатам экспертизы скорректировать содержание и формы заданий.
9. Сформировать выборку для апробационного тестирования.
10. Разработать инструкции для студентов и для преподавателей, проводящих апробацию теста.
11. Провести апробационное тестирование для сбора эмпирических результатов.
12. Выполнить статистическую обработку результатов апробации и наметить пути коррекции теста.
13. Провести редактирование теста: исключить некачественные задания и добавить новые для улучшения характеристик теста (оптимизации трудности заданий, повышения их валидности, улучшения системообразующих свойств заданий, повышения надежности и валидности теста).

14. Повторить этап апробации для доказательства достигнутого повышения качества теста.

2.2 Разработка плана теста и спецификации

Для определения целей тестирования и подхода к составлению теста разрабатывается план теста и его спецификация.

При разработке *плана* определяется:

- процентное содержание дидактических единиц (ДЕ);
- количество заданий по каждой из контролируемых ДЕ;
- формы представления заданий;
- уровень трудности заданий;
- время на выполнение каждого задания;
- время на выполнение теста в целом.

План теста для *промежуточного* контроля должен охватывать знания, умения и навыки по одной или нескольким ДЕ, для *итоговой аттестации* – по всем ДЕ дисциплины в соответствии с программой дисциплины, для проверки *остаточных знаний* по всем ДЕ дисциплины в соответствии с ФГОС по специальности или направлению.

Документом, где содержится информация о целях задачах, плане и структуре тестов является *спецификация теста*. Спецификация очень важна при создании параллельных вариантов, т.к. в ней устанавливается структура содержания всех вариантов одного теста.

Пример спецификации в сокращенном виде приведен в таблице 1. Пример подготовлен по четырем содержательным разделам.

Следует учитывать, что общее количество вопросов или баллов должно быть распределено по темам таким образом, чтобы оно отражало важность, объем и количество учебного времени.

Таблица 1 – Содержательный план теста

№ п/п	Содержательные разделы дисциплины	Планируемые к проверке знания и умения					Суммарное число заданий
		А (10%)	В (20%)	С (30%)	Д (30%)	Е (10%)	
1	I (20%)	1	2	4	4	1	12
2	II (10%)	1	1	2	2	—	6
3	III (30%)	2	4	5	5	2	18
4	IV (40%)	2	5	7	7	3	24
Итого		6	12	18	18	6	60

Примечание: А – знание понятий, определений, терминов; В – знание законов и формул; С – умение применять законы и формулы для решения задач; Д – умение интерпретировать результаты на графиках и схемах; Е – умение проводить оценочные суждения.

В развернутом виде спецификация включает:

- цель, назначение (описание подхода к применению и возможных сфер применения);
- нормативные документы (учебная программа дисциплины), перечень базовых учебников;
- описание контингента испытуемых;
- основные ДЕ и их процентное содержание;
- соотношение заданий по разделам и категориям учебных целей (проверка знаний, умений, навыков или компетенций);
- проверяемые компетенции;
- характеристику полноты охвата требований ФГОС в % (для аттестационных тестов);
- количество ТЗ различных форм;
- рекомендуемое время выполнения ТЗ и теста в целом;
- число параллельных вариантов теста и методику их формирования (параллельность различных вариантов может обеспечиваться за счет подбора определенного количества однотипных, примерно одинаковых по уровню сложности заданий);
- обеспечение валидности, надежности.

Пример спецификации теста в развернутом виде приведен в приложении А.

Следующий этап – разработка заданий, подробно рассмотрен в главе 3.

2.3 Экспертиза тестовых заданий

Экспертиза тестовых материалов позволяет выявить недоработки в ТЗ. Этот этап является обязательным условием для создания аттестационных тестов (рубежного контроля и проверки остаточных знаний студентов). Задания можно считать объективными или корректно сформулированными, если мнения ряда экспертов о назначении задачи, вопроса, о правильности формулировки и пригодности вариантов ответов совпадают между собой. В качестве экспертов тестовых заданий могут быть выступать преподаватели кафедры с большим опытом работы. Возможна внешняя экспертиза специалистами УМО. Если хотя бы один из экспертов высказал замечания к ТЗ, то задача пересматривается, корректируется или исключается.

По окончании экспертизы заполняется заключение эксперта с рекомендациями по коррекции формы и содержания тестовых заданий, приведенное в приложении Б. На основании протокола экспертизы разработчиком проводится заключительная коррекция ТЗ.

2.4 Апробация тестовых заданий

После проведения экспертизы ТЗ, их доработки с учетом результатов экспертизы, проводится апробация заданий. Цель апробации – устранение ошибок и неточностей в заданиях, а также приблизительная оценка трудности теста и ТЗ для улучшения характеристик теста и определение времени, необходимого на решения задания или всего теста целиком.

В зависимости от поставленных задач апробационные исследования организуются как массовые обследования для групп студентов или как небольшие эксперименты над выборкой испытуемых. Желательно использовать большую выборку студентов т.к. уровень объективности результатов апробации напрямую зависит от количества тестируемых. Можно включать интервью студентов, преподавателей и экспертов.

3 ФОРМЫ ТЕСТОВЫХ ЗАДАНИЙ

3.1 Общие требования к ТЗ

В общем виде ТЗ должны:

- соответствовать содержанию дисциплины;
- составляться с учетом соответствующих правил;
- проверяться на практике (апробироваться);
- быть ясными испытуемому.

Вне зависимости от формы ТЗ имеет следующий состав:

1. Инструкция.
2. Текст задания (вопроса).
3. Предполагаемые варианты ответа на задание.
4. Правильный ответ или оценочная схема.

Инструкция должна содержать указания на то, что испытуемый должен сделать, каким образом выполнять задание, в том числе указание на количество верных ответов. Это могут быть фразы:

- «Укажите: правильный ответ; наиболее правильный ответ; все правильные ответы; с точностью до сотых; правильную последовательность; верное утверждение...»
- «Дополните фразу...»
- «Установите соответствие...»
- «Введите с клавиатуры...»

Инструкция должна быть составлена так, чтобы задание и способ его выполнения были абсолютно ясны любому из испытуемых и не приводили к ошибкам. В тестах допускается делать одну инструкцию для группы однотипных заданий, которая помещается в начале теста или данной группы ТЗ в тесте.

Вопрос или *задание* могут быть сформулированы в виде фразы, текста, рисунка, схемы, графика и пр. Написание вопроса начинается со сбора подходящих тестов или с обдумывания ситуаций и тем, по которым можно сгруппировать ряд вопросов. Стиль формулирования задания — не вопрос, а логика высказываний.

Варианты ответов формируются для заданий закрытой формы. Составление тестовых заданий различных форм более подробно рассмотрено в разделах 3.3-3.10.

Для лучшего восприятия задания части ТЗ следует выделить определенным образом, например, вопросительную часть подготовить регистром «ВСЕ ПРОПИСНЫЕ». Во многих тестах ответы кодируются *буквами*, что имеет место преимущественно в естественно-научных дисциплинах, когда сам ответ может быть *численным*.

3.2 Правила оценки тестовых заданий и теста

Одно из важных требований – наличие разработанных правил выставления баллов. Для оценивания результата ответа должно осуществляться основное требование к заданию – ТЗ должно иметь однозначный правильный ответ. Правильный ответ разработчика может заключаться не только в эталонном ответе, но и в описании схемы анализа, вариантов неправильных ответов, считая все остальное правильным.

В большинстве случаев ТЗ оцениваются по номинальной шкале, простейшим случаем которой является дихотомическая шкала (1 балл за правильный ответ, 0 баллов за неправильный ответ). По такой шкале оценивается все задание целиком, а не одна из его частей.

Возможность оценивания правильно выполненной части задания появляется при использовании порядковой (градуированной) шкалы. Баллы при этом выставляются за выбор варианта, выбор соответствия, выбор ранга, выбор дополнения. Такая шкала может быть использована в заданиях с выбором нескольких верных ответов из набора, заданиях на установление правильной последовательности, заданиях на установление соответствия, заданиях открытой формы. Максимальное количество баллов, которое студент может получить за ТЗ, устанавливается разработчиком исходя из сложности задания. Рекомендуется применять схему оценивания не более, чем с тремя оценочными категориями (0, 1, 2). Баллы испытуемому выставляются в зависимости от числа правильно выбранных ответов или полноты ответа (Все верные ответы - 2 балла, 1 ошибка - 1 балл, остальные варианты - 0 баллов).

Итоговая оценка по тесту формируется путем суммирования набранных баллов и отнесения их к общему количеству вопросов в

ТЗ. Помножив полученное значение на 100%, приводим итоговую оценку к традиционной оценке следующим образом:

«2 (неудовлетворительно)» – менее 50%

«3 (удовлетворительно)» – 50% - 70%

«4 (хорошо)» – 70% - 90%

«5 (отлично)» – 90% - 100% .

Специфика перехода на ФГОС ВПО состоит в необходимости проведения оценки компетенций будущих выпускников. Оценивание результатов учебной деятельности можно представить в виде уровней компетентности студентов:

– *Критический уровень*: характеризует низшую степень обученности. Студент демонстрирует недостаточный уровень знаний по данной дисциплине. Это самая низкая ступень овладения студентом знаниями.

– *Репродуктивный уровень*: при оценке результатов обучения студент показывает знание основных понятий, идентифицирует их и воспроизводит, т.е. этот уровень характеризует потенциальные возможности студента.

– *Базовый уровень*: студент воспроизводит ранее усвоенную информацию и решает стандартные задачи, выполняя самостоятельно типовые действия.

– *Продуктивный уровень*: характеризует уровень мастерства, позволяющий студенту гибко подходить к решению более сложных заданий, обосновывать и доказывать правильность выбранного способа их решения.

– *Высокий уровень*: свидетельствует о способности студента творчески (нестандартно) мыслить, использовать имеющуюся систему интегрированных знаний и умений в новой нестандартной ситуации. Студент, показывающий эту наивысшую степень обученности (на данном этапе обучения), способен к обобщению и переносу установленных закономерностей на новые явления.

Для проведения оценки результатов учебных достижений студентов на различных этапах обучения более предпочтительным представляется разработка компетентностно-ориентированных заданий, составленных как кейс-задания или задания открытой формы со свободно-конструируемым ответом с применением градуированной шкалы оценки задания. Например, для оценки результа-

тов выполнения заданий в компетентностном формате можно использовать следующие критерии:

Таблица 2 – Критерии оценки результатов выполнения заданий

Параметр	Уровень освоения компетенции
ситуация разносторонне проанализирована, даны ответы на все вопросы, <i>ошибок нет</i> , предложены обоснованные аргументы и приведены примеры эффективности аналогичных решений, компетенции, осваиваемые в ходе изучения дисциплины, присутствуют в полном объеме	высокий
ситуация разносторонне проанализирована, даны ответы на все вопросы, <i>допущено не более 1 ошибки</i> , предложены обоснованные аргументы и приведены примеры эффективности аналогичных решений, компетенции, осваиваемые в ходе изучения дисциплины, присутствуют в полном объеме	базовый
ситуация поверхностно проанализирована, даны ответы на все вопросы, <i>допущено более 2 ошибок</i> , обоснованных аргументов не предложено	понимания
ситуация практически не проанализирована, ответов на вопросы нет	недостаточный

Наивысший балл в градуированной шкале за компетентностно-ориентированное задание определяется его сложностью и начисляется за задание, выполненное без ошибок. Заметим, что удельный вес оценки за такое задание в ТЗ должен быть достаточно большим, так как компетентностно-ориентированные задания относятся к трудным заданиям высокого уровня сложности.

3.3 Классификация форм ТЗ

ТЗ может быть представлено в одной из четырех стандартизованных форм, как показано на рисунке 2:

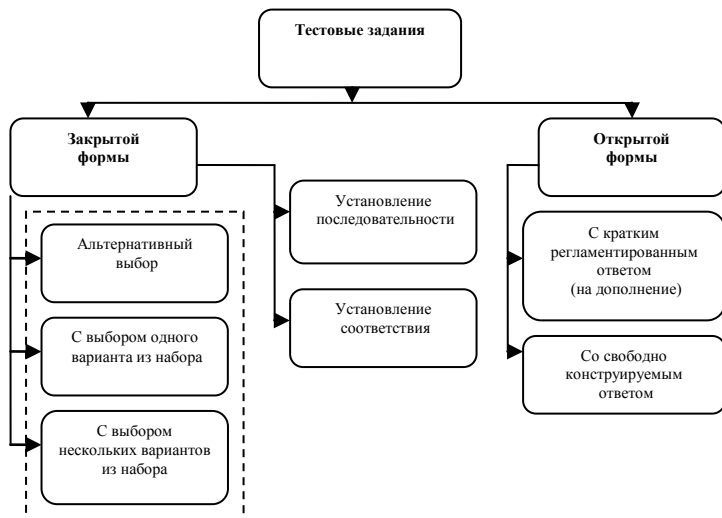


Рисунок 2 – Основные формы тестовых заданий

Выбор формы ТЗ зависит от того, какой вид знаний следует проверить. Для оценки знаний *конкретных фактов, названий, имён, дат, понятий* лучше использовать тестовые задания с выбором ответа или задания с кратким регламентированным ответом на дополнение. *Ассоциативные знания* (знания о взаимосвязи определений и фактов, авторов и их теорий, сущности и явлений, о соотношении между различными предметами, законами, датами) лучше проверять заданиями на установление соответствия, а процессуальные знания (знания правильной последовательности различных действий, процессов) – заданиями закрытой формы на определение правильной последовательности.

Форма предъявления ТЗ влияет на его уровень трудности. Самые легкие – задания закрытого типа с единичным выбором, так как направлены на воспроизведение имеющихся знаний. Самая трудная форма ТЗ - на упорядочение и на соответствие.

Следует учитывать, что задания со свободно-конструируемым ответом используются только при бланковом тестировании ввиду сложности схемы анализа ответа обучаемого. При компьютерном контроле оптимальными являются тестовые задания закрытой формы с выбором ответа, на установление соответствия и на установление правильной последовательности.

3.4 Задания закрытой формы с выбором ответа

3.4.1 Общие требования к тестовым заданиям закрытой формы с выбором ответа

Задания закрытой формы состоят из текста основной части и предложенного множества готовых ответов, среди которых правильными могут быть один или нескольких ответов (множественный выбор). Они наиболее популярны для контрольных тестов ввиду простоты разработки, а также выполнения задания обучаемыми, технологичности в компьютерном тестировании. Недостатком является высокая вероятность угадывания правильного ответа. Скорректировать этот недостаток позволяют качественные дистракторы.

Для ТЗ закрытой формы с выбором ответа имеются два основных типа инструкций «Укажите (отметьте) правильный ответ» и «... наиболее правильный ответ».

Основные требования к ТЗ закрытой формы:

1. Основная часть задания формулируется в форме утверждения.
2. В задании должен содержаться только один вопрос или проблема.
3. Основная часть задания формулируется предельно кратко, ясно и лаконично. Желательно не более одного предложения из семи-восьми слов.
4. Основная смысловая нагрузка должна приходиться на основную часть (длинный вопрос и короткие ответы).
5. Из текста задания исключаются все ассоциации, способствующие выбору правильного ответа с помощью догадки.
6. Задания не должны быть тривиальными.
7. В тексте заданий должна быть устранена всякая двусмысленность или неясность формулировок.

8. Некорректны формулировки вида: «С каким утверждением Вы согласны» и др., содержащие оценочные суждения и выясняющих субъективное мнение обучаемых по какому-либо вопросу.
9. Задание должно иметь правильный ответ. Формулировки «правильного ответа нет», «все перечисленное верно» или «все ответы неправильные» недопустимы.
10. В основной части не следует использовать слова «все», «ни один из», «всегда», «никогда», «и т.д.», «и т.п.»
11. Отрицательные утверждения желательно использовать редко. Нельзя использовать двойное отрицание.
12. Все ответы к одному заданию должны быть приблизительно одной длины либо правильный ответ может быть короче других, но не во всех заданиях теста.
13. Все дистракторы к каждому заданию должны быть равновероятно привлекательными для испытуемых.
14. При формулировке содержательной части нежелательно использовать утверждения, определения из учебника или курса лекций.
15. Не следует формулировать задания-ловушки, присутствующие в психологических тестах (скорее всего в заблуждение будут введены наиболее осведомленные студенты).

Многие из приведенных рекомендаций следует учитывать при создании заданий на установление соответствия и правильной последовательности.

3.4.2 Задания альтернативного выбора

Альтернативная форма задания является самой простой. К каждой задаче дается только два варианта ответов. Испытуемый должен выбрать один из них – «да – нет», «правильно – неправильно» и пр.

Пример 1. ЗНАЧЕНИЕ АБСОЛЮТНОЙ ВЛАЖНОСТИ ОТ ТЕМПЕРАТУРЫ ВОЗДУХА

- A. Зависит
- B. Не зависит

Пример 2. ВОЗДУШНЫЕ ДВИЖЕНИЯ В ЦИКЛОНЕ

А. Восходящие

В. Нисходящие

Использование заданий альтернативных ответов в виде отдельного вопроса, одиночно, как правило, нежелательно для использования в тестировании, но длинные серии таких вопросов имеют определенные преимущества. Поэтому эти задания рекомендуется применять сериями к одному элементу знаний. Это относится к крупным определениям, сложным процессам, к графикам, диаграммам, таблицам, тем элементам знания, которые могут быть структурированы или разбиты на более мелкие части. Например, про сложный объект можно создать серию вопросов, полностью выявляющих его свойства (приложение В). Задание считается выполненным, если на все задания в серии получен верный ответ.

3.4.3 Задания с выбором одного варианта из набора

Задания с выбором одного варианта из набора являются приоритетными для любых контрольных и аттестационных тестов, также полезны для целей обучения. В данной форме возможно представление широкого спектра учебного материала, например расчетных задач, мини-кейсов. В формулировке задания обычно указано: «Укажите правильный ответ».

Пример 3.

Z, м	0	30
T, °C	4.00	2.00
S, ‰	25.50	26.80

В ТАБЛИЦЕ ПРИВЕДЕНА ТЕМПЕРАТУРА И СОЛЕННОСТЬ НА РАЗНЫХ ГЛУБИНАХ. ТОЛЩИНА ОБРАЗОВАНИЯ ЛЬДА ПРИ КОНВЕКЦИИ, ДОХОДЯЩЕЙ ДО ГОРИЗОНТА 30 М, СОСТАВЛЯЕТ:

А. 1м

В. 2м

С. 3м

Д. 6м

Пример 4. СРЕДИ ПЕРЕЧИСЛЕННЫХ АНОМАЛИЙ ВОДЫ ИСКЛЮЧИТЕ ОШИБОЧНУЮ:

- A. температура замерзания
- B. объем
- C. плотность
- D. теплопроводность
- E. электропроводность
- F. диэлектрическая проницаемость

Пример 5. НАИБОЛЬШИЙ ВКЛАД В ДОННЫЕ ОТЛОЖЕНИЯ ВНОСЯТ ... ОСАДКИ:

- A. пирокластические (вулканогенные)
- B. биогенные
- C. терригенные
- D. хемогенные
- E. полигенные

Для заданий с выбором ответа можно воспользоваться некоторыми приемами, рассмотренными в разделе 3.5.

Задания оцениваются по дихотомической шкале (двухбалльная оценка).

3.4.4 Задания с выбором нескольких вариантов из набора

Форма с множественным выбором ответов требует более глубокого анализа имеющихся ответов и дает возможность проверить не только знания, но и умения использовать их в сложных ситуациях. Так же преимуществом этих заданий является хорошая защищенность от угадывания. Основная часть обычно формулируется в виде: «Укажите все правильные ответы».

Пример 6. МЕТОДЫ ВЫДЕЛЕНИЯ ВОДНЫХ МАСС В ОКЕАНЕ:

- A. TS – анализ
- B. Метод сеток
- C. Кластерный анализ
- D. Выделение фронтальной поверхности

Пример 7. УСТОЙЧИВОЙ СТРАТИФИКАЦИИ СУХОГО ВОЗДУХА СООТВЕТСТВУЕТ ИЗМЕНЕНИЕ ТЕМПЕРАТУРЫ ВОЗДУХА С ВЫСОТОЙ:

- A. возрастание или неизменность
- B. убывание, но медленнее, чем предписывает γа
- C. убывание согласно γа
- D. убывание быстрее, чем предписывает γа

Схема оценивания задания с выбором нескольких верных ответов получается более сложной, чем для заданий с выбором одного ответа. В задании такого рода желательно 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).

3.5 Задания на установление соответствия

В заданиях на установление соответствия необходимо найти соответствие (или приравнять части, элементы, понятия) – между элементами двух списков (множеств). Первое множество содержит постановку проблемы, второе включает элементы, подлежащие выбору. Эти задания эффективны при установлении связи между явлениями и закономерностями, между понятиями и их значениями, причинами и следствиями. Они способствуют развитию навыков классификации и полезны для целей обучения.

При составлении заданий на установление соответствия следует руководствоваться следующими правилами:

1. В формулировке задания использовать определяющий термин «соответствие» или производное слово того же корня – «соответствует», «соответствуют» и т.д.
2. Содержание задания формулируется в виде двух множеств с соответствующими заголовками – «задающее» и «множество выбора».
3. Каждому элементу задающего множества должен соответствовать элемент выбора.

4. Элементы каждого множества подбираются по одному основанию. Это может быть текст (имя, термин, понятие, словосочетание), символ (формула, сочетание формул, условный знак, обозначение) или их сочетание.
5. Состав элементов, входящих в множество выбора, должен превышать численный состав задающего множества не менее чем на один элемент.
6. Число элементов одного множества не должно превышать десяти. Если их больше, лучше составить более одного задания.

Сложность этого задания состоит в оценивании ответов. Возможна двухбалльная оценка (0 и 1) за задание целиком, также можно оценивать каждый ответ по отдельности. В связи с этим, задания такого типа стараются не включать в аттестационное тестирование.

Пример 8. УСТАНОВИТЕ СООТВЕТСТВИЕ МЕЖДУ ВИДАМИ И ИСТОЧНИКАМИ ТУРБУЛЕНТНОСТИ:

Задающее множество	Множество выбора
Виды турбулентности	Источники турбулентности
1. Крупномасштабная _____	А. поверхностные волны
2. Среднемасштабная _____	В. внутренние волны
3. Мелкомасштабная _____	С. приливные волны
	Д. синоптические вихри
	Е. волны Росби
	Ф. дрейфовые течения
	Г. волны инерционных колебаний

Пример 9. УСТАНОВИТЕ СООТВЕТСТВИЕ МЕЖДУ ТИПАМИ И ХАРАКТЕРИСТИКАМИ ПЕРЕНОСА В ТРОПОСФЕРЕ УМЕРЕННЫХ ШИРОТ:

Задающее множество	Множество выбора
Типы переноса	Характеристики переноса
Западный (W)	А. Быстро движущие волны малой амплитуды
Восточный (E)	В. Хорошо выраженные высотные гребни и ложбины
Меридиональный (С)	С. Макротурбулентный обмен происходит наиболее интенсивно
	Д. Среднее поля давления характеризуется мощными барическими образованиями

3.6 Задания на установление правильной последовательности

Задания на установление правильной последовательности предназначены для выявления знаний последовательности явлений, процессов, действий, событий и др. Испытуемый должен упорядочить элементы определенного множества в соответствии с заданной закономерностью или заданным правилом, критерием, признаком.

Правила для разработки заданий этого типа:

1. Основная часть обычно формулируется в виде: «Установите правильную последовательность...», «Расположите в правильном порядке...», «Ранжируйте...»
2. Ответы на задания представляются в виде пронумерованного списка в правильном порядке
3. Если в задании подразумевается иерархия, то следует указать с чего начинать отсчет – от младшего или старшего (высшего) элемента.

Пример 10. УСТАНОВИТЕ ПРАВИЛЬНУЮ ПОСЛЕДОВАТЕЛЬНОСТЬ ПОЯВЛЕНИЯ ФОРМ ОБЛАЧНОСТИ ПРИ ПРОХОЖДЕНИИ ТЕПЛОГО ФРОНТА, РАСТАВИВ НОМЕРА В ОТВЕДЕННЫХ МЕСТАХ:

- слоисто-дождевые
- разорванно-слоистые
- высоко-слоистые
- перисто-слоистые
- перистые

В заданиях этой формы также необходимо предусмотреть систему подсчета оценочных баллов в случае, если только часть элементов упорядочена верно.

3.7 Задания открытого типа

Задания открытого типа бывают двух видов, один из которых предполагает получение испытуемым строго регламентированных по содержанию и форме представления кратких правильных ответов (дополнения). Второй – задания со свободно конструируемыми ответами, в которых испытуемые составляют развернутые ответы, произвольные по содержанию и форме представления и содержащие полное решение задачи с пояснениями.

С помощью заданий на дополнение удобно проверять знание названий, фактов, свойств, признаков, дат, причинно-следственных отношений. Для задач на вычисление, задач с формулами в качестве ответа эта форма так же представляется оптимальной.

Основные требования к ТЗ на дополнение:

1. Текст задания должен быть предельно простым по синтаксической конструкции и содержать минимальное количество информации, необходимое для правильного выполнения задания.
2. В задании должно быть только одно дополняемое слово, символ и т.д., место для которого рекомендуется обозначить прочерком или точками (поле ввода при компьютерном тестировании).

3. Прочерк ставится на месте ключевого элемента, знание которого является наиболее существенным для контролируемого материала.
4. Все прочерки в заданиях для одного теста рекомендуется делать равной длины.
5. Место для ответа лучше отводить в конце задания или как можно ближе к концу.

Пример 11. НЕФРОНТАЛЬНАЯ ЛИНИЯ (ИЛИ ПОЛОСА), ВДОЛЬ КОТОРОЙ ИМЕЮТ МЕСТО ПРОЦЕССЫ КОНДЕНСАЦИИ, НАЗЫВАЕТСЯ ЛИНИЕЙ _____

Пример 12. ОТНОСИТЕЛЬНАЯ ВЛАЖНОСТЬ 100%, ТОГДА ВОДЯНОЙ ПАР НАХОДИТСЯ В СОСТОЯНИИ _____

К сложностям при компьютерной реализации данной формы можно отнести необходимость учета синонимов правильного ответа.

Задания оцениваются по дихотомической шкале (1 балл за правильный ответ, 0 баллов за неправильный ответ).

Задания со *свободно конструируемым ответом* наиболее близки к контрольным работам. Они позволяют оценивать сложные учебные достижения. Ответами на такие задания могут быть математические расчеты, анализ кейса, эссе. Трудность заключается в разработке эталона оптимального ответа вместе со стандартизованными правилами оценки результатов его выполнения. Правильность ответа оценивается экспертно, поэтому данную форму заданий следует разрабатывать только в тех случаях, когда не могут быть использованы более простые формы. Эта форма имеет применение только для бланкового тестирования.

Основные рекомендации по разработке заданий со *свободно конструируемым ответом*:

1. Длина и сложность ответа может варьировать в широких пределах. В инструкции желательно ввести ограничения на максимальную длину ответа для каждого задания;
2. Формулировка задания должна включать постановку проблемы, эталон выполнения и оценочные критерии. Постановочная часть должна быть максимально четкой, чтобы устранить воз-

- возможные отклонения в правильных ответах студентов от эталона выполнения;
3. Выбор временных рамок для выполнения каждого задания должен позволять студенту сформулировать достаточно развернутый ответ и успеть его записать.
 4. Такие задания оцениваются большим количеством баллов, чем задания предыдущих форм.

Пример 13. ПЕРЕЧИСЛИТЕ ТИПЫ ВОЗДУШНЫХ МАСС (ВМ) СООТВЕТСТВЕННО ПОЛОЖЕНИЮ ОСНОВНЫХ ТЕРМИЧЕСКИХ ПОЯСОВ. УТОЧНИТЕ ПОЛОЖЕНИЕ ОЧАГА ФОРМИРОВАНИЯ И БАРИЧЕСКУЮ СИСТЕМУ, С КОТОРОЙ СВЯЗАНО ФОРМИРОВАНИЕ ВМ:

Критерии оценивания задания	Балл
Правильный ответ должен содержать следующие элементы: <ol style="list-style-type: none"> 1. Перечислены все типы ВМ. 2. Точно определены очаги формирования и барические системы. 	3 балла
Ответ содержит: <ol style="list-style-type: none"> 1. Перечислены все типы ВМ. 2. Определены очаги формирования. 3. Ошибочные (или отсутствуют) сведения о барических системах (одна и более). 	2 балла
Ответ содержит: <ol style="list-style-type: none"> 1. Перечислены все типы ВМ. 2. Ошибочные (или отсутствуют) сведения об очагах формирования и барических системах (один и более случаев). 	1 балл
Ответ неправильный (все случаи, не соответствующие баллам 1, 2, 3) или отсутствует.	0 баллов

3.8 Сравнительная характеристика форм ТЗ

Выбор оптимальной формы заданий обычно связан со спецификой содержания теста. При этом необходимо учитывать достоинства и недостатки каждой из форм, приведенных в таблице 3, и принимать компромиссное решение.

Таблица 3 – Достоинства и недостатки различных форм ТЗ

Форма	Достоинства	Недостатки
ТЗ закрытой формы с альтернативными ответами	Благодаря краткости позволяют охватить большой объем материала, легко разрабатываются, результаты выполнения обрабатываются быстро и с высокой объективностью	Стимулируют механическое запоминание, поощряют угадывание, требуют увеличения количества заданий и времени тестирования для компенсации эффекта угадывания
ТЗ закрытой формы с выбором вариантов из набора	В силу краткости формулировок позволяют охватить большой объем содержания, обеспечивают возможность автоматизированной проверки, и, соответственно, получение развернутых статистических характеристик ТЗ и высокую объективность оценок	Требуют значительной работы преподавателей при подборе дистракторов, не годятся для проверки продуктивного уровня деятельности и когнитивных умений
ТЗ закрытой формы на установление соответствия	Просты в разработке, идеально подходят для оценивания ассоциативных знаний и проведения текущего контроля, уменьшают эффект угадывания	Громоздки по форме представления, используются для проверки репродуктивного уровня деятельности и алгоритмических умений.
ТЗ закрытой формы на установление правильной последовательности	Подходят для выявления классификационных знаний (построение логических последовательностей, технологических цепочек, алгоритмов, хронологических последовательностей), уменьшают эффект угадывания.	см. ТЗ закрытой формы на установление соответствия
ТЗ открытой формы на дополнение	Просты в разработке, исключено угадывание, частично годятся для автоматизированной проверки	Проверяют в основном знание фактологического материала или понятийного аппарата, иногда приводят к неоднозначным правильным и частично правильным ответам.
ТЗ открытой формы со свободно конструируемыми ответами	Позволяют оценивать сложные учебные достижения, в том числе творческий уровень деятельности, легко формулируются, исключают угадывание	Требуют сложной процедуры проверки, значительного времени выполнения, не позволяют охватить значительный объем содержания предмета, не годятся для автоматизированной проверки.

3.9 Составление заданий с разным принципом композиции

При составлении ТЗ используется ряд логических и методических оснований, который позволяет конструировать содержание.

Принцип *фасетности* позволяет создавать несколько вариантов одного и того же задания, а, следовательно, несколько параллельных вариантов теста, благодаря замене одного (или нескольких) слов (символов, чисел) в базовом задании, которое превращает его в другое (другие), аналогичное по содержанию и трудности. Все элементы фасета должны принадлежать к одной ДЕ. Этот принцип построения ТЗ является универсальным и относится ко всем формам. Задания (как в примере) имеют вид:

Пример 14. ТУМАНЫ {ОХЛАЖДЕНИЯ, ИСПАРЕНИЯ, СМЕШЕНИЯ} К {ФРОНТАЛЬНЫМ, ВНУТРИМАССОВЫМ}

- А. Относятся
- В. Не относятся

Принцип *импликации* отличается от принципа *фасетности* логической формой условного суждения.

Пример 15. ВЫСОТА НИЖНЕЙ ГРАНИЦЫ ОБЛАКОВ {НЕ БОЛЕЕ 100, 100-200, БОЛЕЕ 200} МЕТРОВ. ТОГДА НАЗЕМНЫЕ ОБЪЕКТЫ ОБНАРУЖИВАЮТСЯ С ВЫСОТЫ... МЕТРОВ.

Ответы задания, построенные по принципу *противоречия*, подбираются так, чтобы второй отрицал первый.

Пример 16. ЗНАЧЕНИЕ АБСОЛЮТНОЙ ВЛАЖНОСТИ ОТ ТЕМПЕРАТУРЫ ВОЗДУХА

- А. Зависит
- В. Не зависит

Пример 17. ВОЗДУШНЫЕ ДВИЖЕНИЯ В ЦИКЛОНЕ

- А. Восходящие
- В. Нисходящие

Принцип *противоположности* ответов близок по смыслу к принципу *противоречия*. Отличается тем, что противоположные

ответы допускают возможность существования других переходных понятий и состояний, что увеличивает число вариантов ответов.

Пример 18. АДВЕКЦИЯ – ЭТО ДВИЖЕНИЕ ВОЗДУШНЫХ МАСС:

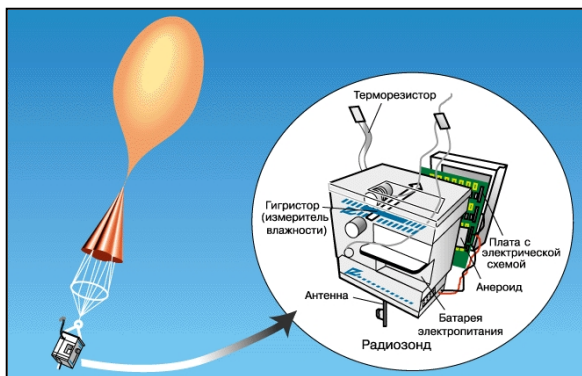
- А. Вертикальное
- В. Горизонтальное
- С. Турбулентное

В следующих примерах *противоположность* вводится в содержание ответа и усиливает содержание задания.

Пример 19. ПРИ КОНВЕКТИВНОЙ НЕУСТОЙЧИВОСТИ ПСЕВДОПОТЕНЦИАЛЬНАЯ ТЕМПЕРАТУРА С ВЫСОТОЙ:

- А. Убывает
- В. Возрастает
- С. Сначала убывает, затем возрастает

Пример 20. ТЕРМОРЕЗИСТОРЫ И ТЕРМИСТОРЫ ЯВЛЯЮТСЯ ДАТЧИКАМИ ТЕМПЕРАТУРЫ. ЕСЛИ ТЕМПЕРАТУРА ВОЗРАСТАЕТ, ТО СОПРОТИВЛЕНИЕ ТЕРМОРЕЗИСТОРОВ:



- А. Возрастает
- В. Уменьшается
- С. Возрастает, затем уменьшается

При использовании принципа *однородности* подбираются такие ответы, которые относятся к одному роду, виду, отображают основные стороны, грани явления. При соблюдении этого принципа обычно создается полная система ответов, в которой нечего добавить. Задания в этом случае воспринимаются испытуемыми наилучшим образом. Данный принцип используется во всех формах закрытого типа и максимально рекомендуется для использования:

Пример 21. УКРУПНЕНИЕ ОБЛАЧНЫХ КАПЕЛЬ – ЭТО:

- A. Коагуляция
- B. Конденсация
- C. Сублимация

Одним из приемов усиления эффективности является использование сходных по написанию или звучанию букв, цифр, знаков, слов или словосочетаний:

Пример 22. ЛИНИИ ОДИНАКОВОГО ЗНАЧЕНИЯ ГЕОПОТЕНЦИАЛЬНЫХ ВЫСОТ:

- A. Изогнеты
- B. Изогипсы
- C. Изотахи
- D. Изобары
- E. Изотенденции

Принцип *кумуляции* вариантов. Каждый последующий вариант вбирает в себя содержание предыдущего ответа. Верным считается самый полный правильный ответ. Для формулировки заданий с использованием этого принципа следует использовать выражение «Выберите наиболее правильный ответ»:

Пример 23. ЯВЛЕНИЕ ГАЛЮ СВЯЗАНО С:

- A. Преломлением
- B. Преломлением и отражением
- C. Преломлением, отражением и дифракцией

Пример 24. ОБЛОЖНЫЕ ОСАДКИ ВЫПАДАЮТ ИЗ ОБЛАКОВ:

- A. Слоисто-дождевых
- B. Высоко-слоистых и слоисто-дождевых
- C. Кучево-дождевых, слоисто-дождевых и высоко-слоистых

Студенты обычно выбирают самый длинный ответ, полагая, что он правильный. Поэтому при разработке заданий, соответствующих этому принципу, нужно подготовить некоторую часть правильных ответов, не самых полных по количеству перечисленных элементов.

Принцип *удвоенного противопоставления* понятий в вариантах ответов заключается в комбинации сочетаний двух независимых признаков, каждый из которых может иметь два варианта выраженности, и применяется в заданиях с четырьмя ответами:

Пример 25. БЕРЕГОВОЙ ЭФФЕКТ НАБЛЮДАЕТСЯ В СЛУЧАЕ:

- A. сходимости линий тока, усилении ветра
- B. расходимости линий тока, ослаблении ветра
- C. сходимости линий тока, ослаблении ветра
- D. расходимости линий тока, усилении ветра

Принцип *сочетания* - ответы komponуются из сочетаний слов по два, реже три, четыре в каждом ответе.

Пример 26. ПРИ ПОЛОЖИТЕЛЬНОЙ ТЕМПЕРАТУРЕ ВОЗДУХА РОСТ КАПЕЛЬ ВОЗМОЖЕН ЗА СЧЕТ:

- A. Конденсации и коагуляции
- B. Конденсации и сублимации
- C. Коагуляции и сублимации

В процессе формирования заданий на основе принципа *сочетания* можно руководствоваться правилом цепочки, в соответствии с которым последний элемент в первом ответе становится первым элементом во втором, последний элемент во втором становится первым в третьем, и т.д.

Пример 27. ВЫБЕРИТЕ НАИБОЛЕЕ ПРАВИЛЬНОЕ УТВЕРЖДЕНИЕ: РЕЛЬЕФ ДНА ДЕЛИТСЯ НА ТРИ ТОПОГРАФИЧЕСКИЕ ЗОНЫ:

- A. подводные окраины материка, океанические котловины, срединно-океанические хребты
- B. срединно-океанические хребты, подводные окраины материка, шельфы
- C. шельфы, океанические котловины, глубоководные желоба
- D. глубоководные желоба, подводные окраины материка, шельфы

Принцип *градуирования* вариантов ответов заключается в том, что применяют упорядоченное использование ответов, выраженных числами или словами, отражающими возрастание (убывание) какого-либо качества или свойства.

Пример 28. ВЫСОТА ТРОПОПАУЗЫ В УМЕРЕННЫХ ШИРОТАХ В ЛЕТНИЙ ПЕРИОД:

- A. 5-8 км
- B. 8-10 км
- C. 10-12 км
- D. 14-16 км
- E. 16-18 км

Принцип *собирающей* формы. Ответ необходимо собрать из логически связанных элементов, каждый из которых подбирается из различных групп, содержащих несколько вариантов. Этот принцип полностью исключает возможность угадывания ответа.

Пример 29. ЧТО ПОНИМАЮТ ПОД АДВЕКЦИЕЙ ТЕПЛА В ОКЕАНЕ. КАКИЕ СВОЙСТВА ОНА ПЕРЕНОСИТ?

- A. перенос в горизонтальном направлении
- B. перенос в вертикальном направлении
- C. вихревое перемещение
- D. температура
- E. соленость
- F. скорость

Распределительная форма задания предусматривает необходимость распределения нескольких ответов по вопросам, связанных общими темами. Представляет большие возможности для проверки глубины знаний. В первой колонке приводятся объекты, которые будут подвергнуты анализу, во второй – одно-двух ступенчатое задание, далее трехступенчатое и т.д. до 8 ступеней. Для примера взято двухступенчатое задание:

Пример 30. РАССМОТРИТЕ ВСЕ ХАРАКТЕРИСТИКИ ГРАВИТАЦИОННОГО И СТАТИСТИЧЕСКОГО ПОЛЯ. УКАЖИТЕ, КАКИЕ ХАРАКТЕРИСТИКИ ЯВЛЯЮТСЯ ОБЩИМИ:

Объекты	Анализирующие качества
Оба поля имеют силовую характеристику	A...величину, направление и точку приложения
Оба поля имеют векторный характер	B...интенсивность напряженности поля
Густота силовых линий характеризует	C...как отношение силы, действующей к величине пробного элемента, помещенного в интересующую точку пространства
Силовая характеристика определяется	D...напряженностью поля

Распределительные задания предоставляют возможность планировать глубину анализа в зависимости от целей тестирования. Количество ответов должно соответствовать числу вопросов.

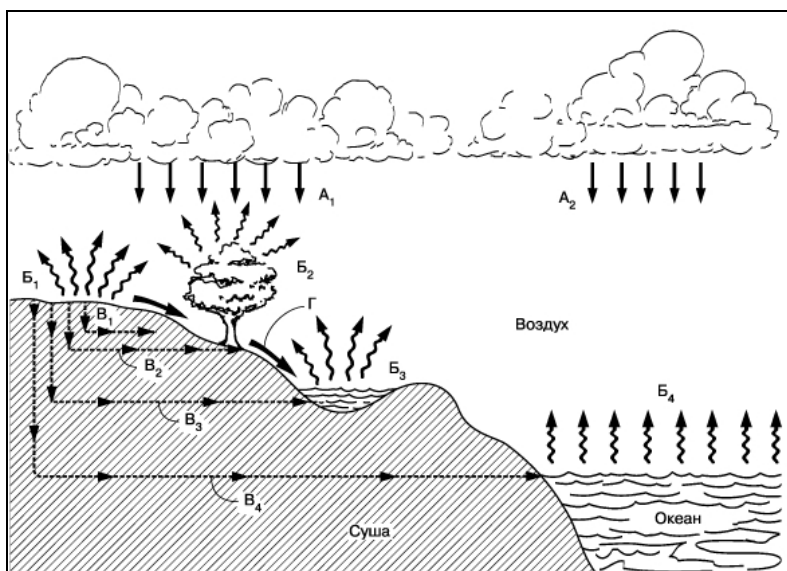
3.10 Кейс-задания

Кейс-задание (специальная проблемная задача) представляет собой задание, состоящее из описания практической ситуации (набора ситуаций) и совокупности вопросов (тестовых задач) к ней. Выполнение студентом кейс-задания требует решения поставленной проблемы в целом и проявления от студентов умения анализировать конкретную информацию, отслеживать причинно-следственные связи, выделять ключевые проблемы, выбирать оптимальные методы решения и др.

Кейс-задание содержит:

- формулировку проблемы;
- подробное описание практической ситуации;
- необходимую для анализа информацию: литературу, иллюстрированный материал, статистические материалы и др.;
- набор заданий в виде комплекса ТЗ различных форм;
- рекомендации по работе с кейсом;
- критерии оценивания результатов (подробно рассмотрены в разделе 3.2 и примере 13).

Пример 31. С акватории Мирового океана, озер, рек постоянно происходит испарение воды, которая конденсируется в атмосфере и, выпадая на поверхность материков и океана, вновь пополняет запасы гидросферы.



КЛИМАТИЧЕСКИЙ ТИП КРУГОВОРОТА ВОДЫ НА ЗЕМЛЕ..
A₁ - осадки, выпадающие над сушей; A₂ - осадки, выпадающие над океаном; B₁ - испарение с суши; B₂ - транспирация растительностью; B₃ - испарение с озер и рек; B₄ - испарение с океана; V₁ - инфильтрация воды в почву; V₂ - потребление воды растительно-

стью; B_3 - подземный сток воды в реки и озера; B_4 - подземный сток воды в океан; Γ - поверхностный сток.

1. Источником всех вод суши являются:

- A_1 - осадки, выпадающие над сушей
- B_3 - подземный сток воды в реки и озера
- B_1 - инфильтрация воды в почву
- B_4 - испарение с океана

2. Озеро может иметь вытекающую реку, если для водосбора верно неравенство:

- $A_1 > B_3$
- $\Gamma + B_3 > B_3$
- $B_2 + B_3 > \Gamma$
- $B_1 + B_3 > A_1$

3. В _____ климатическом поясе круговорот воды происходит интенсивнее, чем в других.

(Введите в поле ответов прилагательное в соответствующем падеже)

Э	К	В	А	Т	О	Р	И	А	Л	Ь	Н	О	М		
---	---	---	---	---	---	---	---	---	---	---	---	---	---	--	--

4. Опишите схему цикла поверхностного стока

океан → атмосфера → суша → поверхностный сток → океан

Как правило, кейсы представляются в печатном виде или на электронных носителях. В настоящее время более популярно мультимедийное представление кейсов, когда в измерительные материалы добавляется видео- и аудио информация.

3.11 Норма трудности тестовых заданий

Для каждого ТЗ следует определить меру трудности, которую необходимо учитывать в процессе контроля знаний. Норма трудности определяется разработчиком тестовых заданий и указывает субъективную величину того, насколько тяжело будет решить данное ТЗ испытуемому с минимальным уровнем подготовки.

Норма трудности зависит от следующих факторов:

– количества концептов (формула, правило, аксиома и т.д.), необходимых для поиска правильного решения ТЗ. Чем больше шагов нужно выполнить для получения правильного ответа, тем выше норма трудности, тем сложнее считается ТЗ;

– формы ТЗ;

– уровня значимости ТЗ;

– глубины расположения ТЗ в спецификации;

– количества дистракторов в заданиях закрытой формы.

При определении трудности задания можно руководствоваться следующим:

Простые задания:

1. «Опознание» какого-то объекта или проверка «знания-знакомства».
2. Выбор одного варианта ответа из многих с помощью знания всего одного концепта.
3. Выявление знания определения односложного базового термина с помощью ТЗ открытого типа.
4. Принадлежность ТЗ основному материалу: ТЗ раскрывает базовое понятие.
5. Нижний уровень иерархии спецификации теста (например, некоторое "Понятие").

Средний уровень:

1. Применение усвоенных ранее знаний в типовых ситуациях (т.е. в тех ситуациях, с которыми знаком испытуемый) или проверка «знаний воспроизведения копии».
2. Средний уровень иерархии спецификации теста (например, "Тема" или "Подтема").

Сложный уровень:

1. Применение усвоенных знаний и умений в нестандартных условиях (т.е. в условиях, ранее не знакомых испытуемому) или проверка «знаний умения и применения».
2. Принадлежность ТЗ дополнительному материалу.
3. Верхний уровень иерархии спецификации теста (например, "Раздел" или "Глава").

Следует отметить, что нарушение правил составления тестовых вопросов, рассмотренных в разделе 3, приводит к появлению ошибок и делает ТЗ непригодными в тестировании. Примеры типичных ошибок в заданиях приведены в приложении Г.

4 АНАЛИЗ КАЧЕСТВА СОСТАВЛЕНИЯ ТЕСТОВ И ТЕСТОВЫХ ЗАДАНИЙ

Статистические расчеты проводятся с целью анализа качества теста целиком и каждого ТЗ с точки зрения его вклада в решаемую тестом задачу педагогического измерения. Эта процедура рекомендуется для тестов всех типов, которые будут использоваться для контроля достижений студентов, а для аттестационных тестов является обязательной.

Для получения количественных статистических характеристик теста и тестовых заданий необходимо выполнить ряд операций.

Общие операции:

- формирование матрицы ответов по результатам тестирования;
- определение индивидуальных баллов и количество ответов.

Расчет качественных характеристик теста:

- среднеарифметического, дисперсии, стандартного отклонения;
- надежности, валидности, гомогенности, оптимальной длины.

Определение качественных характеристик тестовых заданий:

- трудности;
- валидности;
- дискриминативности;
- качества дистракторов.

4.1 Общие операции

Результаты проведения тестирования вносятся в таблицу ответов (таблица 4).

Таблица 4 – Матрица ответов испытуемых

Номера испытуемых, i	Номера заданий						Индивидуальные баллы, X_i
	1	2	...	i	...	n	
1							
2							
...							
N							
Число правильных ответов, R_j							

Подсчитывается индивидуальный балл испытуемого и количество правильных ответов на каждое задание:

$$X_i = \sum_j X_{ij}, \quad (1)$$

$$R_j = \sum_i X_{ij}. \quad (2)$$

4.2 Расчет качественных характеристик теста

Вычисляется среднеарифметическое значение (3) для совокупности индивидуальных баллов:

$$\bar{X} = \frac{\sum_{i=1}^N X_i}{N}. \quad (3)$$

Отклонение балла от среднего $X_i - \bar{X}$ показывает место испытуемого по отношению к среднему баллу.

Для оценивания степени неоднородности результатов тестирования рассчитывается *дисперсия* (4) и *стандартное отклонение* (5). Дисперсия измеряет разброс баллов, полученных испытуемыми. Низкое значение дисперсии указывает на плохое качество теста, поскольку не обеспечивается высокий дифференцирующий эффект. Высокая дисперсия характеризует случай, когда все студенты отличаются по числу выполненных заданий, также требует переработки теста из-за существенного отличия вида распределения баллов от планируемой нормальной кривой.

$$S_x^2 = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N - 1}, \quad (4)$$

$$S_x = \sqrt{S_x^2}. \quad (5)$$

Для анализа отклонения распределения баллов от нормального вида вычисляют значение $\bar{X} - 3S_x$. Значительные отклонения от нуля свидетельствует об отличии от нормального вида распределения.

По формуле

$$\frac{\sum_{i=1}^N (X_i - \bar{X})^3}{NS_x^3} \quad (6)$$

рассчитывается *асимметрия*, которая характеризует степень отклонения распределения баллов от симметрии, значение ее должно приближаться к нулю, отрицательное значение показывает, что тест труден, большинство студентов получили баллы ниже среднего.

Для графического представления по полученным результатам строится гистограмма распределения частоты индивидуальных баллов, представленная на рисунке 3.



Рисунок 3 – Частотное распределение сгруппированных индивидуальных баллов

Обычно считается, что хороший тест обеспечивает нормальное распределение индивидуальных баллов, как показано на ри-

сунке 4, когда среднее значение находится в центре распределения, а остальные концентрируются вокруг среднего значения по нормальному закону ($\approx 70\%$ в центре, значение $\bar{X} - 3Sx$ близко к 0).

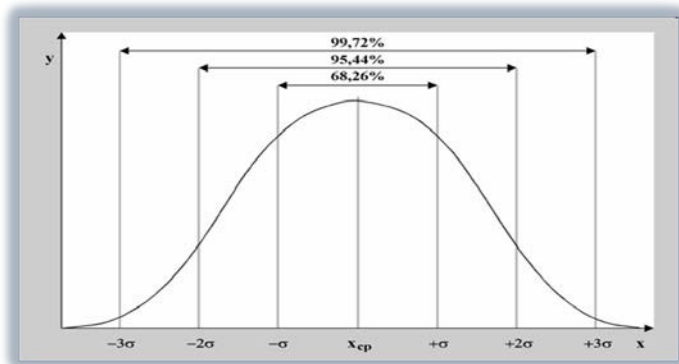


Рисунок 4 – Нормальная кривая распределения индивидуальных баллов

Оценка *надежности* теста (7) строится на подсчете корреляции между двумя наборами данных (X и Y) одного и того же теста, значение которых измеряются по прошествии некоторого времени.

$$(r_{tt})_{\text{рет}} = \frac{N \sum_{i=1}^N X_i Y_i - \left(\sum_{i=1}^N X_i \right) \left(\sum_{i=1}^N Y_i \right)}{\sqrt{N \sum_{i=1}^N (X_i)^2 - \left(\sum_{i=1}^N X_i \right)^2} \sqrt{N \sum_{i=1}^N (Y_i)^2 - \left(\sum_{i=1}^N Y_i \right)^2}}, \quad (7)$$

где $(r_{tt})_{\text{рет}}$ – коэффициент надежности теста по ретестовому (двукратному) методу;

X_i – индивидуальный балл i -го испытуемого в первом тестировании,

Y_i – индивидуальный балл i -го испытуемого во втором тестировании

Приемлемый коэффициент надежности должен превышать значение 0,69.

Выделяют виды *валидности*:

– содержательная, которая дает ответы на вопросы, охватывает ли содержание теста весь комплекс программных требований

к значению данного конкретного предмета и насколько данные задания (выбранные из множества возможных) пригодны для оценки знаний по этому предмету. Содержательная валидность определяется с помощью экспертной оценки всех ТЗ.

– критериальная – корреляция результатов тестирования с внешним к тесту критерием.

– прогностическая – корреляция с результатами других педагогических измерений, предназначенных для этих же целей.

Гомогенность (однородность, предметная чистота заданий) теста оценивается по преобразованному коэффициенту корреляции Пирсона:

$$\varphi_{jl} = \frac{p_{jl} - p_j p_l}{\sqrt{p_j q_j \cdot p_l q_l}}, \quad (8)$$

где p_{jl} – доля испытуемых, выполнивших правильно оба задания с номерами j ;

p_j – доля испытуемых, правильно выполнивших j -е задание;

p_l – доля испытуемых, правильно выполнивших l -е задание теста;

$q_j = 1 - p_j$ и $q_l = 1 - p_l$, соответственно, – доля испытуемых неправильно выполнивших j и l задание теста. Значения от 0 до 0,3 считаются оптимальными.

Для определения *оптимальной длины теста* руководствуются следующим: длина теста должна быть такой, чтобы вероятность неправильной аттестации испытуемых, имеющих индивидуальный балл выше (или ниже) на 10% критериального балла, характеризующего неудовлетворительную оценку, не превышала 20%.

4.3 Определение качественных характеристик тестовых заданий

После проведения пробного тестирования для каждого задания высчитывается *коэффициент трудности* (p_j) по формуле:

$$p_j = \left(1 - \frac{R_j}{N}\right) * 100, \quad (9)$$

где N – общее количество испытуемых, R_j – количество испытуемых, правильно ответивших на данный вопрос.

К примеру, если ТЗ выполнило менее 20% испытуемых, оно слишком сложное, а если более 80% - слишком легкое (значения 80 и 20 % являются условными). Преобладание слишком легких или трудных заданий увеличивает ошибку измерения знаний.

По результатам апробации часть заданий будет отбракована. Исходя из этого, для апробации, следует взять некоторый запас заданий. Количество заданий должно превышать итоговое не менее чем в 1,5 раза.

Для оценки *валидности* применяют коэффициент точечно-биссерной корреляции между результатами выполнения каждого задания и тестовым баллом испытуемых:

$$(r_{pbis}) = \frac{(\bar{X}_2)_j - (\bar{X}_1)_j}{S_x} \sqrt{\frac{(N_2)_j (N_1)_j}{N(N-1)}}, \quad (10)$$

где $(\bar{X}_2)_j$ – среднее значение индивидуальных баллов, выполнивших j задание верно;

$(\bar{X}_1)_j$ – среднее значение индивидуальных баллов, выполнивших j задание неверно;

S_x – стандартное отклонение;

N_2 – количество испытуемых, выполнивших j задание верно;

N_1 – количество испытуемых, выполнивших j задание неверно;

N – общее количество испытуемых.

Задание имеет валидность, если $r_{pbis} \approx 0,5$. Задания с отрицательными r_{pbis} следует удалять из теста, т.к. их выполняют слабые испытуемые и не выполняют сильные.

Дискриминативность (дифференцирующая способность) определяет, насколько хорошо ТЗ дифференцирует лучших и слабых студентов. Наиболее простым и наглядным показателем является индекс дискриминации задания, который рассчитывается по результатам тестирования двух «контрастных» групп (обычно 27% «лучших» и 27 «слабых» студентов):

$$(r_{дис})_j = (p_1)_j - (p_2)_j, \quad (11)$$

где p_1 и p_2 – доли студентов в лучшей и слабой группе, ответивших на j задание правильно.

Если $r_{\text{дис}}$ близко к 1, то данное задание обладает высокой дифференцирующей способностью. Показатель $r_{\text{дис}} > 0,3$ можно считать удовлетворительным.

Для проведения статистического *анализа качества дистракторов* проводится расчет процента испытуемых из выборки, которые выбрали тот или иной дистрактор из выборки:

- дистракторы, которые выбираются малым количеством студентов (например, менее 5%), должны быть удалены из теста.

- если какие-то дистракторы выбираются большим количеством студентов, то их так же стоит пересмотреть;

- когда доля испытуемых, выбравших правильный ответ значительно меньше, чем доля испытуемых, выбравших некий дистрактор, то следует проанализировать содержание.

ЗАКЛЮЧЕНИЕ

Новые требования к освоению образовательных программ обуславливают применение различных технологий и форм контроля достижений результатов обучения. Однако, проектирование тестов, как части оценочных средств вуза, является достаточно сложной, многокомпонентной задачей. В оценочные средства должны быть включены задания различных форм, в том числе и компетентностно-ориентированные тесты для приближения обучающихся к их профессиональной деятельности. Проблема состоит и в том, что лишь немногие преподаватели владеют теорией педагогических измерений. В дальнейшем предстоит:

1. повышение квалификации в области теории и практики тестовых технологий педагогического состава;
2. создание оптимальной системы средств контроля для текущей, промежуточной и др. аттестаций;
3. совмещение исторически сформированных методов контроля с новыми инновационными моделями оценки знаний, умений, навыков (ЗУН) и компетенций;
4. внедрение разнообразных форм тестового контроля, позволяющего осуществлять непрерывный мониторинг формирования ЗУН и компетенций;
5. развитие методического, программного и технологического обеспечения тестирования;
6. создание банка тестовых заданий по различным дисциплинам.

Решение данных задач позволит достигнуть такого уровня контроля и управления качеством образования, который обеспечит беспрепятственное признание квалификации выпускников университета мировыми образовательными системами.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Федеральный государственный образовательный стандарт высшего профессионального образования по направлению подготовки 280400 прикладная гидрометеорология [Электронный ресурс]. URL: <http://quality.rshu.ru/content/111/pgm>.
2. Григорьева Е.Г., Трубина М.А., Черемных А.В. Проблемы и решения компьютерного тестирования // Ученые записки РГГМУ. – 2010. – № 14. – С. 187-198.
3. Звонников В.И., Чельшкова М.Б. Современные средства оценивания результатов обучения: Уч. пос. 4-е изд. – М.: Издательский центр «Академия», 2009. – 223 с.
4. Аванесов В.С. Композиция готовых заданий.: Уч. пос. – М.: Центр тестирования, 2002. – 238 с.
5. Ефремова Н.Ф. Современные тестовые технологии в образовании. - М., Логос, 2003. – 176 с.
6. Майоров А.Н. Теория и практика создания тестов для системы образования. – М.: Интеллект-Центр, 2002. – 296 с.
7. Караушев В.Ф., Терентьева Л.В., Тягунова Т.Н. Проектирование банка программно-дидактических тестовых заданий / под ред. Васильева В.И. – М.: Издательство ИКАР, 2005. – 63 с.
8. Переверзев В.Ю. Критериально-ориентированное педагогическое тестирование в профессиональном образовании (методология, теория, практика): Монография. М.: Федеральный институт развития, 2008. - 248 с.
9. Кабанова Т.А., Новиков В.А. Тестирование в современном образовании: Уч. пос. – М.: Высшая школа, 2010. – 381 с.
10. Крокер Л., Алгина Дж. Введение в классическую и современную теорию тестов. – М.: Логос, 2010. – 664 с.
11. Григорьева Е.Г., Гагулина Н.Л., Черемных А.В. Разработка тестов и тестовых заданий: Учебно-методическое пособие. – СПб: Изд-во РГГМУ, 2012. – 64 с.

ПРИЛОЖЕНИЕ А. ПРИМЕРЫ СПЕЦИФИКАЦИИ ТЕСТА ПО ДИСЦИПЛИНЕ «ОБЩАЯ ОКЕАНОЛОГИЯ»

Таблица А1 – Краткая спецификация (по ДЕ и категориям учебных целей)

Наименование ДЕ	А	В	С	D	Е	Кол-во заданий
Общие сведения о мировом океане. Понятие о пространственно-временной изменчивости океанологических процессов.	3	1	2	0	2	8
Перемешивание и обмен в океане	1	1	0	1	1	4
Термика океана. Взаимодействие океана и атмосферы.	2	1	1	0	1	5
Водные массы и фронты Мирового океана. Тонкая структура гидрофизических процессов в океане.	1	1	0	1	1	4
Итого	7	4	3	2	5	21

Примечание: А – знание понятий, определений, терминов; В – знание законов и формул; С – умение применять законы и формулы для решения задач; D – умение интерпретировать результаты на графиках и схемах; Е – умение проводить оценочные суждения.

Подробная спецификация теста

1. Назначение. Тест предназначен для промежуточной аттестации студентов 2 курса океанологического факультета Российского государственного гидрометеорологического университета по направлению подготовки 280400 – Прикладная гидрометеорология, профиль подготовки – Прикладная океанология, обучающихся на 3 семестре.

2. Спецификация составлена в соответствии с требованиями ФГОС по направлению Прикладная гидрометеорология, профиль подготовки – Прикладная океанология.

3. Тест составлен в рамках критериально-ориентированного подхода.

4. Курс «Общая океанология».

5. Количество вариантов теста – 5.

6. Количество заданий в тесте – 30.

7. Количество заданий, контролирующих:

- уровень знаний, определений, законов, формул - 40%
- умений, навыков их применения - 60%.

8. Код проверяемых компетенций – ОК-1, ПК-1, ПК-9

9. Количество форм заданий – 5, из них:

- задания с выбором правильного ответа из набора –10;
- задания с выбором нескольких правильных ответов из набора – 7;
- задания на дополнение – 3;
- задания со свободно конструируемым ответом – 5;
- задания на установление соответствия – 5.

10. Время, отведенное на выполнение каждого задания:

- с выбором одного правильного ответа из набора – 1 минута;
- с выбором нескольких правильных ответов – 3 минуты;
- с кратким дополняемым ответом – 2 минуты;
- установления соответствия – 2 минуты;
- со свободно конструируемым ответом - 5 минут.

Общее время – 80 минут.

11. Для создания тестов, предъявления их испытуемым было проведено бланковое тестирование, использована система тестирования MOODLE, система тестирований РГГМУ.

12. Экспертное заключение кафедры по содержанию тестов прилагается.

13. Исправления по замечанию экспертов внесены в тестовые задания для предъявления испытуемым.

Таблица А2 – Количественное и процентное соотношение заданий по разделам

№	Контролируемая ДЕ	Количество заданий в % от общего кол-ва заданий	Номера заданий
1	Общие сведения о мировом океане. Понятие о пространственно-временной изменчивости океанологических процессов	4 (13%)	1-4
2	Перемешивание и обмен в океане	4 (13 %)	5, 8, 11-12
3	Термика океана. Взаимодействие океана и атмосферы	5 (17 %)	6-7, 9-10
4	Водные массы и фронты Мирового океана. Тонкая структура гидрофизических процессов в океане	5 (17 %)	13-17
5	Оптические и акустические свойства морских вод	6 (20 %)	18-19, 27-30
6	Химические и биологические процессы в океане. Природные ресурсы Мирового океана	6 (20 %)	20-26

ПРИЛОЖЕНИЕ Б. ПРОЕКТ ЭКСПЕРТНОГО ЗАКЛЮЧЕНИЯ
ЭКСПЕРТИЗЫ ТЕСТОВЫХ ЗАДАНИЙ

ЭКСПЕРТИЗА ТЕСТОВЫХ ЗАДАНИЙ

Дата _____

Тест _____

(дисциплина)

Эксперт _____

(Ф.,И.,О.)

Количество тестовых заданий в тесте: _____

Специальность: _____

Тестовые задания (ТЗ), прошедшие экспертизу (обвести номер ТЗ)

№ 1	№ 2	№ 3	№ 4	№ 5	№ 6	№ 7	№ 8	№ 9	№ 10	№ 11	№ 12	№ 13	№ 14
№ 15	№ 16	№ 17	№ 18	№ 19	№ 20	№ 21	№ 22	№ 23	№ 24	№ 25	№ 26	№ 27	№ 28

Оценить качество ТЗ по критериям:

(зачеркнуть несоответствие)

1. Соответствие целям тестирования (содержание ФГОС):

№ 1	№ 2	№ 3	№ 4	№ 5	№ 6	№ 7	№ 8	№ 9	№ 10	№ 11	№ 12	№ 13	№ 14
№ 15	№ 16	№ 17	№ 18	№ 19	№ 20	№ 21	№ 22	№ 23	№ 24	№ 25	№ 26	№ 27	№ 28

2. Лексическая и стилистическая грамотность

№ 1	№ 2	№ 3	№ 4	№ 5	№ 6	№ 7	№ 8	№ 9	№ 10	№ 11	№ 12	№ 13	№ 14
№ 15	№ 16	№ 17	№ 18	№ 19	№ 20	№ 21	№ 22	№ 23	№ 24	№ 25	№ 26	№ 27	№ 28

3. Пригодность вариантов ответов

№ 1	№ 2	№ 3	№ 4	№ 5	№ 6	№ 7	№ 8	№ 9	№ 10	№ 11	№ 12	№ 13	№ 14
№ 15	№ 16	№ 17	№ 18	№ 19	№ 20	№ 21	№ 22	№ 23	№ 24	№ 25	№ 26	№ 27	№ 28

КОРРЕКЦИЯ:

Стилистические ошибки и описки:

№№ ТЗ:

Изменение формы:

№№ ТЗ:

Изменение сложности:

№№ ТЗ:

Ошибка в правильном варианте ответа:

№№ ТЗ:

Замечания и предложения эксперта:

Заключение: Тест соответствует (не соответствует) требованиям, предъявляемым к программно-дидактическим тестовым материалам (ПДТМ) и может быть использован в образовательном процессе РГГМУ
(зачеркнуть ненужное)

ПРИЛОЖЕНИЕ В. ПРИМЕР СЕРИИ АЛЬТЕРНАТИВНЫХ ВОПРОСОВ

УРАВНЕНИЕ РАДИАЦИОННОГО БАЛАНСА
ДЕЯТЕЛЬНОГО СЛОЯ

$$B=(S' +D)_{\times}(1-A)-E_{\text{эф}}$$

- (1). При расчете инсоляции постоянная Стефана-Больцмана – *используется/не используется.*
- (2). Интегральный коэффициент прозрачности характеристической инсоляции – *является/не является.*
- (3). При уменьшении оптической массы атмосферы значение инсоляции ослабится – *верно/неверно.*
- (4). Рассеянная радиация закону Буге – *подчиняется/не подчиняется.*
- (5). Количество облаков на значение рассеянной радиации – *влияет/не влияет.*
- (6). При уменьшении альбеда происходит увеличение поглощаемой радиации – *верно/неверно.*
- (7). С возрастанием влажности почвы альбеда увеличится – *верно/неверно.*
- (8). Альбеда суточным ходом – *обладает/не обладает.*
- (9). При понижении влажности деятельного слоя значение эффективного излучения увеличивается – *верно/неверно.*
- (10). Расходная часть радиационного баланса состоит только из $E_{\text{эф}}$ – *верно/неверно.*
- (11). При переходе к пасмурному небу в ночное время радиационный баланс – *уменьшится/увеличится.*

ПРИЛОЖЕНИЕ Г. ПРИМЕРЫ ОШИБОК В ЗАДАНИЯХ

Пример	Недостатки
<p>Что такое коэффициент тепловой инерции термометра?</p> <p>А. Это безразмерный коэффициент, показывающий отношение пределов измерения термометра.</p> <p>В. Это время, в течение которого разность температур между датчиком термометра и окружающей средой уменьшается в e раз.</p> <p>С. Это коэффициент, показывающий отношение выходного и входного параметра термометра.</p> <p>Д. Это та температура, до которой термометр должен нагреться или охладиться.</p>	<p>– избыточная информация в ответах;</p> <p>– вопросительная форма задания</p>
<p>Метод Г.Н. Милейко используется для прогноза:</p> <p>А. температуры воды в теплый период года</p> <p>В. температуры воды в холодный период года</p> <p>С. температуры воды в глубоких слоях океана</p> <p>Д. температуры воды в условиях ветро-волного перемешивания</p>	<p>– некорректные формулировки ответов (перегруженность одинаковыми словами)</p>
<p>Управление контроллером Unilog с помощью кнопок: как посмотреть показания различных каналов контроллера? Выберите два варианта, каждый из которых составляет часть решения.</p> <p>А. Двукратное нажатие кнопки [ON], когда контроллер выключен</p> <p>В. Однократное нажатие кнопки [ON], когда отображается измеренное значение</p> <p>С. Нажатие кнопки [ON] и удержание более 5 с, когда контроллер включен</p> <p>Д. Нажатие кнопки [ON] и удержание более 5 с, когда контроллер выключен</p> <p>Е. Нажатие кнопок [+]/[-]</p>	<p>– задание с избыточной информацией в основной части</p>
<p>Какой вид информации наиболее важен для выработки проектных решений:</p> <p>А. Климатическая</p> <p>В. Метеорологическая</p> <p>С. Синоптическая</p> <p>Д. Гидрологическая</p>	<p>– недостаточная информация в формулировке основной части;</p> <p>– вопросительная форма задания</p>

Пример	Недостатки
<p>Почему погибают леса?</p> <p>A. Потому что их уничтожают пожары</p> <p>B. Потому что это естественный процесс</p> <p>C. Потому что их уничтожает человек</p> <p>D. Потому что это часть общей деградации и разрушения биоты</p> <p>E. Потому что меняется химический состав атмосферы и почвы</p>	<p>– см. недостатки предыдущего задания;</p> <p>– избыточная информация в ответах</p>
<p>ИЗМЕРИТЕЛЬНАЯ ЧАСТЬ ГИДРОЛОГИИ</p> <p>_____ (гидрометрия)</p>	<p>– очевидность решения</p>
<p>Сколько типов устойчивости экосистем вы знаете?</p> <p>A. Два типа</p> <p>B. Три типа</p> <p>C. Четыре типа</p> <p>D. Пять типов</p> <p>E. Шесть типов</p>	<p>– задания, в которых испытуемому предлагается высказать мнение;</p> <p>– избыточная информация в ответах</p>
<p>Какие датчики используются в гидрологическом стенде?</p> <p>A. Гидростатический</p> <p>B. Актинометрический</p> <p>C. Доплеровский</p> <p>D. Радиологический</p> <p>E. Барботажный</p> <p>F. Логометрический</p> <p>G. Радарный</p> <p>H. Индукционный измеритель скорости течения</p> <p>I. Поплавковый</p> <p>J. Анемометр</p> <p>K. Характериограф</p> <p>L. Доплеровский профилограф</p> <p>M. Осадкомерный</p> <p>N. Гидрограф</p> <p>O. Ультразвуковой расходомер</p> <p>P. Радиолокационный</p>	<p>– задание содержит очень большое количество дистракторов;</p> <p>– принципиально разные основания в их подборе;</p> <p>– вопросительная форма задания</p>
<p>Укажите факторы, обуславливающие влагоперенос в атмосфере и в почве:</p> <p>A. Гравитационные силы</p> <p>B. Гидрофизические характеристики</p> <p>C. Влажность</p>	<p>– задания, описывающие разные свойства или качества объекта;</p> <p>– для выбора дистрак-</p>

Пример	Недостатки
D. Термический режим	торов используются принципиально разные явления
<p>Как бороться с наркобизнесом?</p> <p>A. Всех причастных расстреливать</p> <p>B. Обеспечить новые экономические условия</p> <p>C. Вести просветительскую работу</p> <p>D. Улучшить медицинское обслуживание</p> <p>E. Применять все имеющиеся в распоряжении общества методы и средства</p>	<p>– некорректная формулировка отдельных ответов;</p> <p>– правильный ответ заметно длиннее остальных;</p> <p>– вопросительная форма задания</p>
<p>Сколько каналов в общей сложности способен поддерживать изучаемый контроллер?</p> <p>A. 1</p> <p>B. 2</p> <p>C. 8</p> <p>D. 32</p>	– присутствуют ненужные слова «в общей сложности», «изучаемый»
<p>Гидростатический датчик фиксируется в трубе с помощью</p> <p>A. Резьбового соединения</p> <p>B. Уплотнительных манжет</p> <p>C. Герметика</p> <p>D. Не фиксируется</p>	– последний ответ противоречит смыслу основной части
<p>Объем флэш-памяти контроллера составляет</p> <p>A. 1 Мб</p> <p>B. 2 Мб</p> <p>C. 4 Мб</p> <p>D. 0,5 Мб</p>	– ответы не упорядочены
<p>Что называют экологическими факторами?</p> <p>A. События</p> <p>B. Явления</p> <p>C. Эффекты</p> <p>D. Причины</p> <p>E. Совокупность вещей, явлений и эффектов, влияющих на состояние экосистемы и человека</p>	<p>– задание с неработающими дистракторами (правильный ответ определяется по принципу наибольшей полноты);</p> <p>– вопросительная форма задания</p>

Пример	Недостатки
<p>Что является предметом исследования агрометеорологии ?</p> <p>A. Погода и растения</p> <p>B. Климат и растения</p> <p>C. Взаимодействие погоды и климата с растениями</p> <p>D. Почвенные условия и растения</p>	<p>– см. недостатки предыдущего задания</p>
<p>Сколько ячеек в диапазоне B2:E5?</p> <p>A. 16</p> <p>B. 10</p> <p>C. 12</p> <p>D. 17</p>	<p>– содержит невалидные задания (тест - информатика, а расчет – математика)</p>
<p>Установите соответствие между элементами списков:</p> <p>A. наука о подземных водах</p> <p>B. наука о движениях земной коры</p> <p>C. раздел геодезии</p> <p>D. наука о землетрясениях</p>	<p>– в вопросах на установление соответствия используются неоднородные объекты</p>
<p>Кто и когда ввел в экологию термин "биогеоценоз"?</p> <p>A. Тьер де Шарден, 1905</p> <p>B. Д. Холлинг, 1958</p> <p>C. Ф. Валери, 1960</p> <p>D. В. Жадин, 1931</p> <p>E. А. Тэнсли, 1935</p>	<p>– зависимость результата от памяти или от других индивидуальных особенностей испытуемого, а не от уровня развития тех умений и навыков, для оценки которых разрабатывается тест</p>