

## ОЦЕНКА КАЧЕСТВА ЭКЗАМЕНАЦИОННЫХ ЗАДАНИЙ ЭКСПЕРТНЫМ МЕТОДОМ

Н.Ф. Стась, В.В. Мамонтов, А.И. Галанов

Томский политехнический университет  
E-mail: stanif@mail.ru

*На примере химии проведена комплексная оценка тестовых заданий, разработанных для контроля качества подготовки студентов технических университетов. Определено их соответствие перечню и уровню контролируемых знаний, тестологическим требованиям, современной терминологии и символике дисциплины, трудность и другие характеристики. Выявлены задания, не пригодные для контроля и измерения знаний. Указаны направления работы по оптимизации качества заданий.*

Основой любого способа контроля знаний и умений являются задания, которые выполняют студенты. От их соответствия целям и содержанию обучения зависят результаты контроля и меры по совершенствованию учебного процесса. В Томском политехническом университете экзамен и рубежный контроль по химии студентов технических (нехимических) направлений и специальностей проводится по тестовой технологии [1]. Контроль с помощью тестов более объективен, по сравнению с классической формой контроля, по трем причинам: 1) отсутствуют субъективные факторы, влияющие на оценку; 2) ответ экзаменуемого сравнивается с эталоном; 3) тест является объемным инструментом контроля, т. к. проверяются знания и умения по всем темам дисциплины, а не её отдельные фрагменты. Но результаты тестового контроля чувствительны к качеству используемых заданий, т. е. если недостатки задания в традиционном экзаменационном билете могут быть компенсированы дополнительными вопросами при устном собеседовании, то при тестировании это невозможно.

Объективная оценка качества заданий производится при обработке результатов их использования на основе математических моделей современной тестовой теории. Но обязательным этапом разработки тестов является предварительная экспертная оценка заданий [2]. Практика тестирования (как в России, так и за рубежом) свидетельствует о том, что высокое качество тестовых заданий не могут гарантировать даже самые опытные преподаватели. Поэтому задания, предназначенные для целей контроля и измерения знаний, должны проходить строгую объективную экспертизу. Оценка заданий даже одним экспертом дает гарантии устранения целого ряда недостатков [3], а наиболее оптимальное число экспертов в группе – не менее 3-х человек [4].

В данном исследовании проведена экспертиза заданий по 4-м темам нового набора (банка) заданий, который создаётся по плану совершенствования и развития применяемой технологии, это: 1) химическая связь и строение вещества, 2) основы химической термодинамики, 3) окислительно-восстановительные реакции, 4) электрохимические процессы.

Экспертиза заданий обычно проводится на их соответствие тестологическим требованиям. В дан-

ной работе сделана попытка реализации комплексного подхода к оценке качества заданий. Для каждого задания мы определяли 10 характеристик: 1) соответствие перечню контролируемых знаний и умений, 2) уровень контролируемых знаний и умений, 3) форма задания (согласно классификации по форме ответа), 4) тип расчётной задачи, 5) взаимосвязь подзаданий в двойных заданиях, 6) трудность, 7) значимость, 8) соответствие тестологическим требованиям, 9) соответствие современному содержанию, терминологии и символике дисциплины, 10) пригодность к применению.

**1. Соответствие перечню контролируемых знаний и умений.** Организуя работу по формированию банка заданий, кафедра пришла к выводу о необходимости иметь согласованный между всеми преподавателями перечень целей изучения по всем темам и элементам содержания дисциплины. При составлении перечня использовались: 1) описание предметного содержания дисциплины в ГОС ВПО и образовательном стандарте университета; 2) типовая программа изучения дисциплины, разработанная учебно-методическим объединением; 3) рабочие программы лекторов, прошедшие аттестацию и аккредитацию; 4) содержание учебников и учебных пособий, имеющих рекомендательный гриф; 5) опыт преподавания дисциплины на кафедре.

В окончательном варианте перечня контролируемые элементы содержания дисциплины сгруппированы по 12 темам, которые изучаются студентами всех общетехнических направлений. Это фундаментальный материал, на основе которого изучаются специальные разделы химии, соответствующие требованиям направлений. Фактически это рабочая программа дисциплины в целевой форме, в которой по каждому элементу содержания определены виды деятельности студента: знать, объяснить, вычислять, устанавливать, записывать и т. д. В качестве примера приведем перечень элементов содержания по теме «Основы химической термодинамики».

1. Знать предмет химической термодинамики и её значение, знать параметры стандартного состояния, объяснять понятие «химическая система» и знать типы систем (открытая, закрытая, изолированная, неизолированная).
2. Знать классификацию реакций в химической термодинамике (экзотермические и эндотерми-

- ческие, изохорные и изобарные, самопроизвольные и несамопроизвольные).
3. Объяснять понятия внутренней энергии и энтальпии системы, знать их взаимосвязь (первый закон термодинамики).
  4. Определять, не проводя вычислений, тип реакции (экзотермическая или эндотермическая) при известном и неизвестном значении энтальпии этой реакции.
  5. Знать определение понятия «стандартная энтальпия образования вещества», уметь вычислять её, если известны масса (объем) простого вещества, взаимодействующего с другим простым веществом и количество выделяющегося при этом тепла.
  6. По значению энтальпии образования вещества вычислять количество тепла, которое выделяется или поглощается при получении любой массы (объема – для газа) этого вещества из простых веществ.
  7. Уметь записывать термохимические уравнения реакций, знать их особенности.
  8. Вычислять энтальпию реакции по основному закону термохимии посредством алгебраических преобразований термохимических уравнений.
  9. Вычислять энтальпию реакции по следствию основного закона термохимии.
  10. Вычислять количество тепла, которое выделяется или поглощается при получении известной массы (объема – для газа) продукта или при израсходовании известной массы или объема (для газов) реагента.
  11. Объяснять физико-химический смысл энтропии системы и стандартной энтропии вещества, определять по уравнению реакции, не проводя расчетов, как изменяется энтропия (увеличивается, уменьшается) при протекании реакции.
  12. Вычислять энтропию реакции и, по результату вычисления, определять направление протекания реакции в изолированной системе (второй закон термодинамики).
  13. Вычислять энергию Гиббса химической реакции при стандартной и нестандартной температуре и делать выводы о возможности и направлении её самопроизвольного протекания.
  14. Сравнить термодинамическую устойчивость соединений по значениям их стандартной энергии Гиббса образования.

При разработке и внедрении тестовых технологий контроля знаний обсуждается проблема: надо ли проверять знание всего материала или выделять для контроля самые важные («ключевые») элементы содержания. Имеются сторонники как первого, так и второго подхода. По нашему мнению, отказ от контроля усвоения какой-то части дисциплины может привести к тому, что многие студенты будут игнорировать соответствующий материал, и их

знания по предмету будут отрывочными. Составители и эксперты заданий должны исходить из того, что содержательная валидность набора заданий по каждой теме и по всей дисциплине должна быть полной, а распределение заданий по элементам содержания – равномерным.

**2. Уровень контролируемых знаний и умений.** В процессе обучения у студентов формируются определенные виды познавательной деятельности, которые соответствуют достижению некоторого уровня обученности и которые могут быть установлены контролирующими заданиями. Экспертиза заданий по этому показателю проводилась с опорой на широко известную четырехуровневую классификацию достижений обучаемых, разработанную В.П. Беспалько [5].

**3. Форма задания.** В классической тестологии [6] известно четыре формы заданий: 1) закрытые (выбор правильного ответа среди дистракторов), 2) открытые (ответ формулируется самостоятельно), 3) установление соответствия, 4) выявление последовательности. В химии имеется более детальная классификация, в которой используется 10 форм заданий [7]:

1. Выбрать один ответ, «замаскированный» неверными, отвлекающими ответами-дистракторами (ВО).
2. Выбрать несколько ответов (ВНО).
3. Выбрать дополнение к основной части задания (ВД).
4. Написать самостоятельно пропущенное слово, число, символ и т. п. (НС).
5. Заполнить пробел в таблице, схеме, рисунке (ЗП).
6. Оценить суждения (ОС).
7. Установить соответствие (УС).
8. Установить последовательность (УП).
9. Определить стехиометрические коэффициенты в уравнении реакции (ОК).
10. Решить расчетную задачу (РЗ).

Задания, в которых студент решает расчетную задачу (РЗ), относятся к третьему уровню их учебных достижений. Упражнения, в которых необходимо самостоятельно написать ответ (НС, ЗП, ОК, УС и УП) – это, как правило, задания второго уровня. Заданиями на выбор правильного ответа, «замаскированного» дистракторами (ВО, ВД, ВНО и ОС), обычно проверяются знания и умения первого уровня.

**4. Тип задания.** Этот показатель относится только к расчетным задачам, которые в химии подразделяются на прямые, обратные, комбинированные и с межпредметными связями [8]. К прямым относятся задачи, решаемые подстановкой величин, заданных в условии задачи, в известную формулу, которая является математическим выражением закона, правила или определения. В обратных задачах

искомая величина находится в правой части формулы, и её необходимо сначала выразить с помощью математических преобразований. Комбинированные задачи объединяют материал нескольких тем дисциплины, а задачи с межпредметными связями отражают связи данной дисциплины с другими. В тестологии существует запрет на использование комбинированных задач, поскольку они нарушают однородность теста. Но химия – часть естествознания, где химические явления и закономерности тесно связаны с физическими, а математика является инструментом их описания. Поэтому, если студент знает, например, уравнение состояния газа или закон действующих масс, но не умеет проводить по ним расчеты, то цена такого знания (по классификации В.П. Беспалько это первый-второй уровни) не может быть высокой.

**5. Взаимосвязь подзаданий.** Взаимосвязь подзаданий – особая оценка, обусловленная тем, что нами используются, в основном, двойные задания: студент на каждое задание должен привести два ответа. При этом два ответа могут быть независимы, последовательны и зависимы. Независимыми ответы бывают в том случае, когда проверяется знание различных закономерностей или явлений. Ответы последовательны, если первый является промежуточным, а второй окончательным результатом выполнения задания. Ответы зависимы, когда один из них является следствием другого. Следует иметь в виду, что заданий с зависимыми ответами не должно быть, т. к. они, в действительности, являются заданиями с одним ответом. Что касается заданий с последовательными ответами, то их можно использовать, при этом они могут служить индикатором их самостоятельного выполнения: если промежуточный ответ неправильный, а окончательный – верный, то он получен студентом в результате подсказки или использования шпаргалки.

**6. Трудность.** Трудность задания является субъективной оценкой эксперта. Поэтому детальную классификацию и точную оценку заданий по уровню трудности провести невозможно. По трудности задания подразделялись нами на три категории: легкое, средней трудности и трудное. Оценивая трудность задания, мы принимали во внимание проработанность данного элемента содержания на лекциях и в учебниках, восприимчивость студентов к данному материалу, выполнялись ли подобные задания на практических занятиях, имеются ли похожие задания в индивидуальном домашнем задании, тип задачи (например, обратная задача, как правило, труднее прямой), число умственных действий, которые необходимо выполнить, чтобы получить ответ, как выполняются подобные задания на рубежном контроле и т. д. Следует иметь в виду, что трудность зависит не от формы задания, а от его содержания.

**7. Значимость.** Под значимостью задания понимается его важность (значение) с точки зрения общих целей изучения дисциплины. Этот показатель не имеет точного определения и количественной

оценки, поэтому определяется экспертами (высокая значимость, средняя, низкая), исходя из своего опыта и своих субъективных оценок. Например, по теме «Основы химической термодинамики», по которой выше приведены элементы её содержания, значимость соответствующих заданий оценивается преподавателями и экспертами по-разному. Одни считают наиболее важным знание закономерностей термохимии, другие – законов (начал) термодинамики, а третьи – классификаций систем и реакций в химической термодинамике, т. к. они, по их мнению, определяют всю структуру знаний студентов по этой теме. В оценках значимости заданий обычно наблюдается заметный разброс, поэтому здесь необходимо сравнение оценок не менее трех экспертов.

**8. Соответствие тестологическим требованиям.** Основная часть и дистракторы в закрытых тестовых заданиях должны удовлетворять определенным специфическим требованиям. В мировой практике такие требования подробно изложены в работе [9].

**9. Соответствие современному содержанию, терминологии и символике дисциплины.** В заданиях недопустимо использование материалов, не соответствующих современному содержанию химии: формул несуществующих соединений; надуманных (не соответствующих справочным данным) характеристик соединений, реакций и растворов; устаревшей символики и терминологии. Такие задания необходимо не только указывать, но и отмечать, что именно в каждом из них не соответствует содержанию дисциплины.

**10. Пригодность.** Вывод о пригодности задания следует из предыдущих характеристик и может быть одним из трех: 1) пригодно, 2) пригодно, но требуется переработка, 3) непригодно и должно быть заменено.

Комплексная оценка проведена нами для 392 заданий, относящихся к четырём указанным выше темам. Установлено, что все задания соответствуют современному содержанию дисциплины и в них используется современная терминология и символика, а тестологическим требованиям соответствуют 383 задания (97,7 %).

Показатели других характеристик, к сожалению, не столь однозначны. Установлено, что самыми лучшими характеристиками обладает набор заданий по теме «Окислительно-восстановительные реакции». По этой теме заданиями охвачены все элементы её содержания. Большинство заданий (80 %) проверяется третий, т. е. высокий, уровень знаний и умений (применение). Значимость заданий высокая (78 %) и средняя (22 %). Ответы во всех двойных заданиях независимы. По форме ответа большинство заданий (89 %) относятся ко второй группе, в которых студент должен самостоятельно определиться с ответом. Недостатком заданий является то, что среди них нет трудных заданий: по оценкам экспертов они все относятся к лёгким (35 %) или обладают средней трудностью (65 %).

По теме «Основы химической термодинамики»

уровень проверяемых знаний и умений в большинстве заданий (83 %) высокий (применение), который обеспечивается использованием расчётных задач. Значимость заданий, в основном (84 %), высокая. Среди заданий имеются как лёгкие (18 %), так и трудные (15 %) задания, но большей частью они отнесены экспертами к заданиям средней трудности (67 %). Вместе с тем, по этой теме экспертизой установлен неполный охват заданиями всех элементов её содержания (имеются задания по 11 элементам из 14) и неравномерное распределение имеющихся заданий по элементам её содержания. Обнаружен и другой недостаток: имеются двойные задания с зависимыми ответами; таких заданий немного (18 %), но их не должно быть вообще. Наконец, большинство расчётных задач (58 из 68, т. е. 85 %) – прямые, что тоже можно рассматривать как недостаток.

По теме «Электрохимические процессы» ответы в двойных заданиях независимые или последовательные, зависимые – только в одном задании. Значимость заданий, в основном (85 %), высокая. Но по другим характеристикам в наборе заданий по этой теме имеются недоработки. Например, на один из элементов содержания (приводить примеры и описывать применение электролиза в химической промышленности, в металлургии и при нанесении покрытий) нет ни одного задания. По остальным семи элементам содержания задания имеются, но распределены они очень неравномерно: разброс от 3 до 29 заданий. Имеется много заданий (16 %), которыми проверяется самый низкий уровень знаний – узнавание.

По теме «Химическая связь и строение вещества» распределение заданий по трудности лучше, чем в других темах: имеются задания лёгкие (18,1 %), средней трудности (60,6 %) и трудные (21,3 %). В двойных заданиях ответы, в основном (91 %), независимые. С другой стороны, экспертами не выявила заданий по двум (из 13) элементам содержания данной темы, а задания, которые соответствуют 11-ти элементам содержания, распределены по ним неравномерно. Кроме того, имеется значительное число заданий (18,1 %), направленных на проверку первого уровня проверяемых знаний и умений (узнавание).

По трём темам (основы химической термодинамики, электрохимические процессы, химическая связь и строение вещества) имеются задания, которые выходят за рамки согласованного перечня элементов содержания. Это тот материал химии, который изучается студентами химических направлений и специальностей, но студентами нехимических направлений его изучение не обязательно. Составителям заданий следует ещё раз проанализировать требования Государственного образовательного стандарта и программу дисциплины, рекомендованную УМО для общетехнических направлений, и согласовать с ними перечень элементов содержания.

Общие результаты экспертизы по всем четырём темам приведены в таблице. Из таблицы видно, что

соответствие заданий перечню контролируемых знаний и умений по данным четырём темам составляет 92,6 %. Это высокий показатель, но соответствие должно быть полным. Этот недостаток необходимо устранять составлением дополнительных заданий при сокращении числа заданий по тем элементам содержания, по которым имеется их явный избыток.

**Таблица.** Характеристики тестовых заданий

Характеристики заданий	Доля заданий, %
Соответствие перечню контролируемых знаний и умений:	
– соответствует	92,6
– не соответствует	7,4
Уровень контролируемых знаний и умений:	
– узнавание	10,2
– воспроизведение	31,4
– применение	58,4
– перенос	–
Форма (классификация по форме ответа):	
– первая группа (РЗ)	26,0
– вторая группа (НС, ЗП, ОК, УС, УП)	40,6
– третья группа (ВО, ВД, ВНО, ОС)	33,4
Тип (для расчётных задач):	
– прямая	84,3
– обратная	13,7
– комбинированная	2,0
– с межпредметными связями	–
Взаимосвязь подзаданий:	
– независимые	62,3
– последовательные	23,4
– зависимые	7,3
Оценка трудности задания:	
– лёгкое	25,8
– средней трудности	65,3
– трудное	8,9
Значимость:	
– высокая	67,3
– средняя	30,4
– низкая	2,3
Соответствие тестологическим требованиям:	
– соответствует	97,7
– не соответствует	2,3
Соответствие содержанию дисциплины:	
– соответствует	100
– не соответствует	–
Вывод о пригодности:	
– пригодны без переработки	84,2
– пригодны, но требуется переработка	11,7
– непригодны и должны быть заменены	4,1

Более чем половиной заданий (58,4 %) проверяется третий уровень знаний и умений (применение), но имеются и такие задания, которыми контролируется самый низкий уровень (узнавание). Этот недостаток следует устранить в первую очередь.

В комплекте имеется 102 (26 %) расчётные задачи. Их небольшое число можно оправдать тем, что по двум темам (химическая связь и строение вещества и окислительно-восстановительные реакции) расчётные задачи в курсе химии технических университетов программой не предусмотрены, но по двум другим темам расчётных задач могло бы быть больше. Следует отметить, что большинство имею-

щихся расчётных задач – прямые (84,3 %) или обратные (13,7 %). Практически нет задач комбинированных и с межпредметными связями, что свидетельствует об отрыве изучаемого материала от реальных проблем, решаемых с помощью химии. Этот недостаток можно преодолеть с помощью учебного пособия [10], которое посвящено этому вопросу.

При составлении двойных заданий важно, чтобы подзадания не были зависимыми (такие задания в действительности являются заданиями с одним ответом). Этому требованию соответствует только 62,3 % заданий. Наличие 23,4 % заданий с последовательными ответами не является недостатком, но 14 заданий (7,3 %) с зависимыми ответами должны быть заменены.

Распределение заданий по трудности среди данных 392 заданий неравномерное: 25,8 % (лёгкие), 65,3 % (средней трудности) и 8,9 % (трудные). Это неправильно: распределение должно быть равномерным.

При определении значимости заданий расхождение оценок экспертов наблюдается наиболее ча-

сто. Усреднённые данные свидетельствуют о том, что 67,3 % заданий оценено экспертами как высокозначимые с точки зрения общих целей изучения дисциплины, 30,4 % обладают средней значимостью и 2,3 % низкой.

Итогом экспертизы является оценка пригодности заданий. Пригодными являются 84,2 % заданий, а переработка требуется для 11,7 % заданий. Признано непригодными и должно быть заменено 16 (4,1 %) заданий.

Подробные результаты экспертизы по каждой теме переданы комиссии, созданной из наиболее опытных преподавателей, перед которыми поставлена задача устранения недостатков, установленных экспертами.

*Выражаем благодарность доцентам кафедры общей и неорганической химии ТПУ, которые приняли участие в экспертизе заданий по отдельным темам: Г.В. Кашкан (химическая связь и строение вещества), Л.Н. Шиян (электрохимические процессы) и Г.А. Вороновой (основы химической термодинамики).*

#### СПИСОК ЛИТЕРАТУРЫ

1. Минин М.Г., Стась Н.Ф., Жидкова Е.В., Родкевич О.Б. Тестовая технология контроля знаний по химии // Известия Томского политехнического университета. – 2005. – Т. 308. – № 4. – С. 281–235.
2. Чельшкова М.Б. Теория и практика конструирования педагогических тестов. – М.: Логос, 2002. – 432 с.
3. Майоров А.Н. Теория и практика создания тестов для системы образования. – М.: Народное образование, 2000. – 352 с.
4. Михайлова Н.С., Минин М.Г. Моделирование экспертизы разрабатываемого дидактического теста // Известия Томского политехнического университета. – 2006. – Т. 309. – № 6. – С. 247–251.
5. Беспалько В.П. Слагаемые педагогической технологии. – М.: Педагогика, 1989. – 190 с.
6. Аванесов В.С. Форма тестовых заданий. – М.: Центр тестирования, 2005. – 156 с.

7. Стась Н.Ф. Классификация и составление параллельных заданий для тестов по химии // Вопросы тестирования в образовании. – 2004. – № 11. – С. 46–52.
8. Стась Н.Ф. Классификация расчетных задач по химии // Развитие творческого мышления в образовательном процессе: Матер. регион. научно-практ. конф. 15-16 февраля 1997 г. / Департамент образования администрации Томской области. – Томск: Изд-во ТГУ, 1997. – С. 26.
9. Евтюхин Н.В., Бондарева Т.В., Дубинина Т.В., Сурыгина И.Ю. Современное состояние методов тестирования знаний и умений за рубежом и в России // Инновации в образовании. – 2004. – № 1. – С. 27–47.
10. Икрин В.М., Стась Н.Ф. Межпредметные связи химии. – Томск: Изд-во ТПУ, 2004. – 102 с.

*Поступила 29.12.2006 г.*