

ОЦЕНКА КАЧЕСТВА ТЕСТОВЫХ ЗАДАНИЙ

Аннотация. В статье рассматривается необходимость разработки методики объективной оценки трудности и сложности тестовых заданий. Проводится дифференциация понятий «сложность» и «трудность».

Ключевые слова: тестовое задание, трудность, факторы сложности.

Необходимость оценки профессиональных квалификаций имеет место на всех стадиях образовательного процесса: от входной аттестации, определяющей базовый уровень подготовки обучающихся 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 государственных образовательных стандартов в практику высшего профессионального образования повышает актуальность оценки профессиональных квалификаций. Практически во всех федеральных государственных образовательных стандартах определено: «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». Тесты, как часть оценочных средств, являются эффективным средством контроля результатов обучения на уровне теоретических знаний, понимания и умения применять знания на практике. Они позволяют не только получить объективную информацию о качестве знаний и умений обучающихся, но и определить наиболее слабо усвоенные разделы, темы, отдельные вопросы и своевременно скорректировать процесс обучения.

Чтобы тесты и тестовые задания могли быть использованы в процессе оценки качества профессионального образования наиболее эффективно, необходимо четко сформулировать и соблюдать общие требования к их составу и содержанию, выработать правила оценки тестов и тестовых заданий.

Современная тестология признает, что качество тестирования зависит от многих параметров, но основной фактор – качество разработки тестов, их валидность. Тестовые задания, удовлетворяющие «внешним» формальным требованиям (правильность формулировки, логическая непротиворечивость, адекватная форма и др.) не являются всегда качественными, позволяющими рассматривать такие задания априори или апостериори как эффективное средство оценки уровня подготовки обучаемых. А. М. Сохор отмечает, что «подлинная активизация познавательной деятельности обучающихся возможна лишь при целесообразной последовательности предъявления задач, а также при учете особенностей тех элементов задач, которые должны стать центром мыслительной деятельности» [6]. Поэтому при построении системы учебных заданий нужно ориентироваться на их когнитивные характеристики, которые являются компонентами сложности. Для создания такой системы педагогу требуется оценить сложность заданий априори, т. е. до того, как они будут выполняться обучающимися.

В классической теории тестирования в качестве характеристики задания принимается его «статистическая трудность» [1].

Большинство исследователей полагает, что трудность тестового задания — это субъективная характеристика, которая зависит от огромного числа «внутренних параметров» чело-

века и не может быть эффективно формализована. «Трудный – это значит требующий большого умственного напряжения, усилий, затруднительный, мудреный». Это определение дает один из критериев оценки трудности тестового задания, а именно - сколько времени потребует испытуемому, чтобы усилия, направленные на поиск правильного ответа были потрачены не зря. Понятие трудности может основываться на статистических оценках, например, чем меньше правильных ответов, тем труднее задание. Однако Р. А. Гильманов и др. считают, что «трудность, как и сложность, могут выступать объективными, а точнее, объективированными характеристиками, поскольку зависят от вполне объективных алгоритмов человеческого мозга» [3].

Какие факторы влияют на трудность тестового задания. В классической теории тестов трудность задания исчислялась как отношение верно данных на него ответов к общему количеству испытуемых. Естественным образом, чем ближе полученный результат к 1 (единице), тем большее количество испытуемых с ним справилось, следовательно, тем легче задание. Поэтому со временем показатель трудности задания стали рассчитывать как отношение неверно данных ответов к общему количеству испытуемых. Суть от этого, правда, поменялась не очень сильно. Формула позволяет констатировать, что с заданием X испытуемые справляются лучше (или хуже), но не дает никакого представления о том, почему. Соотношение верно и неверно выполненных заданий является ключевым элементом в вычислении вероятности правильного ответа, предложенном Г. Рашем. Он предложил «опираться только на «наблюдаемые параметры», такие как количество верных ответов, данных на ряд заданий испытуемым N, и количество верных ответов, данных всеми испытуемыми на задание X» [8]. Оба примера пригодны для обработки результатов тестирования. Если их использовать на различных этапах текущего контроля, они станут эффективным инструментом коррекции дальнейших действий педагога. Кроме того, расчет вероятности правильного ответа (например, по Рашу) позволит дифференцировать и индивидуализировать процесс обучения.

Многими педагогами и психологами отмечается объективность такой характеристики задания, как его сложность.

«Сложный – это значит состоящий из нескольких частей или элементов, образованный посредством соединения, сложения частей». Очевидно, из этого определения становится ясным как можно установить сложность тестового задания. Для этого достаточно проанализировать количество элементов знаний, охватывающих это тестовое задание, а также установить насколько глубоко испытуемый при ответе должен знать предметную область, чтобы правильно ответить на это ТЗ. Другими словами, сложность может быть связана с количеством умственных операций, которые нужно совершить, чтобы достичь правильного ответа. Если испытуемый не знает некоторые из этих операций - то для него задание будет сложным, если знает - более легким.

Часто степень трудности учебного задания не совпадает с его сложностью. Например, А.Н.Захаров и А.М.Матюшкин отмечают, что степень трудности учебного задания не совпадает с его сложностью. Степень сложности учебного материала характеризуется реальной (объективной) насыщенностью учебного задания и формой его изложения, а степень трудности всегда предполагает соотнесение подлежащего усвоению учебного материала с ранее усвоенным учебным материалом и интеллектуальными возможностями обучающихся. Трудность учебной задачи часто объясняется тем, что учащиеся часто не знают тех операций, которые надо производить, чтобы найти решение. Если систему операций для решения некоторого класса задач назвать методом решения, то, по его мнению, трудность связана с незнанием метода, с незнанием, как нужно думать в процессе решения, как и в какой последовательности надо действовать с условиями задачи. Возникающие затруднения объясняются тем, что педагог часто старается дать знания о содержании изучаемого и значительно меньше заботится о том, как надо думать, рассуждать. Подобное истолкование пересекается с идеей о связи сложности задания с числом операций, которые необходимо совершить для достижения успеха. Эти определения трудности и сложности являются, по большей части, психологическими; они полезны при психологическом анализе содержания тестовых заданий.

В общем случае сложность и трудность определяются разработчиком ТЗ и указывают субъективную величину того, насколько тяжело будет решить данное тестовое задание испытуемому с минимальным уровнем подготовки за определенное время. Можно различать два вида сложности и трудности - теоретическая (априорная) и фактическая (апостериорная). Априорная сложность и трудность определяются экспертами до проведения тестовых испытаний. В процессе апробации банка тестовых заданий и расчета по определенным методикам получают апостериорные сложность и трудность.

Нужно учитывать, что при использовании банка тестовых заданий при тестировании трудность и сложность может адаптироваться в соответствии с аудиторией, т.е. при получении статистических данных по ответам на задания можно выявить насколько легко задание воспринимается тестируемыми, и какие заключения на него даются. В соответствии с этим устанавливаются фактическая (апостериорная) трудность и сложность тестового задания.

«Важнейшее назначение сложности и трудности ТЗ - использование при алгоритмах адаптивного тестирования. При отсутствии информации о сложности и трудности невозможна адаптация тестовых заданий по текущему уровню знаний испытуемого. Более того, при ошибочном задании трудности и сложности алгоритмы адаптивного тестирования будут работать некорректно, и в результате будет выставлена оценка уровня учебных достижений, имеющая большую погрешность» [4].

Отмечая различия между понятиями «сложность» и «трудность», исследователи подчеркивают связь между ними в виде корреляции. При планировании контрольных мероприятий в том случае, если трудность и сложность заданий коррелируют между собой с высоким коэффициентом, появляется возможность «путем оценки сложности конкретных задач того или иного типа прогнозировать их трудность, обходясь тем самым без непосредственного измерения последней, которое, как правило, значительно более трудоемко» [2].

Следовательно, знание сложности заданий позволяет осуществить предсказание трудности их решения обучающимися.

Таким образом, актуальной является необходимость разработки методики объективной оценки сложности заданий. Критерием ее удачности может служить высокая корреляция найденных с ее помощью оценок со значениями статистической трудности, полученными на больших выборках испытуемых. Если такая корреляция будет подтверждена, это послужит основанием для признания возможности применения предложенной методики оценки сложности при построении систем заданий.

Первым шагом оценки сложности должно стать построение измерительной шкалы. Наиболее часто встречается порядковая шкала с тремя градациями сложности заданий: легкое, средней сложности и сложное. По мнению некоторых авторов, тестовые задания также должны быть разделены на три степени сложности: уровень А – задания, рассчитанные на усвоение основных понятий, на простое отображение материала, на уровне узнаваемости и воспроизведения; уровень Б – задания, требующие размышления, охватывающие малый материал, выявляющие умения применять знания в стандартных ситуациях; уровень В – задания, предусматривающие творческое использование приобретенных знаний и позволяющие выявить умения, применять знания в нестандартных ситуациях.

Следующим за определением шкалы шагом оценки сложности является «установка процедуры, в соответствии с которой сложность конкретного задания может быть отнесена к той или иной градации, учитывая, что сложность может определяться несколькими факторами, например, количеством и характером мыслительных операций, необходимых для выполнения задания» [7]. Эти факторы также должны быть оценены по количественной шкале, после чего по ним строится интегральный показатель, характеризующий сложность задания в целом. Интегральные показатели могут определяться следующими способами:

1) обычным суммированием (или усреднением составляющих показателей) – в этом случае все факторы считаются равноправными;

2) суммированием с весовыми множителями (нахождение весового среднего значения), которые отражают значимость фактора. Например, в работе Л. М. Фридмана «каждой группе

операций назначается свой весовой коэффициент, который прямо пропорционален затраченному на выполнение операций времени, тогда сложность задачи рассчитывается как сумма коэффициентов сложности всех выполняемых в ходе решения операций» [7].

В обоих случаях требуется обоснование равнозначности факторов или их статистических весов. Заключительным шагом должно стать отнесение сложности по ее интегральному показателю к той или иной градации шкалы на основании установленных заранее критериев.

Таким образом, сложность – объективная многофакторная количественная характеристика учебного задания, отражающая число и характер мыслительных операций, необходимых для его решения нормативным способом.

При анализе допускаемых обучающимися ошибок рядом авторов выделены следующие факторы сложности заданий:

1) техническая сложность (ТС) – количество действий, которые необходимы для выполнения задания при нормативном алгоритме решения, она зависит от количества аналитических выражений (формул), используемых при решении, и определяет объем «технической» работы и время ее выполнения;

2) когнитивная сложность (КС) включает факторы, связанные с характером мыслительной деятельности, необходимой для выполнения задания (запоминание, знание, воспроизведение, применение и др.);

3) дополнительная сложность (ДС) – фактор, связанный с использованием при решении положений и формул из нескольких разделов дисциплины.

Интегральный показатель сложности S равен сумме трех групп факторов с равными весами: $S = TC + KC + DC$.

Таким образом, количественная оценка сложности учебных заданий, необходимая для прогнозирования трудности решения задач обучающимися, а также упорядочение на ее основе учебного материала, возможна. Она требует обоснованного выделения факторов сложности и вычисления их весов на основе объемных статистических данных о результатах педагогических измерений. Критерием полноты учета факторов может служить корреляция между статистической трудностью задания и его сложностью.

Библиографический список

1. Аванесов В.С. Аванесов В.С. Композиция тестовых заданий. М.: Центр тестирования, 2002. 237 с.
2. Балл Г.А. Теория учебных задач: психолого-педагогический аспект. М.: Педагогика, 1990. 184 с.
3. Гильманов, Р. А. Проблемы дидактики, трудности учебных упражнений/Р. А. Гильманов. — Казань, 1989. — 179 с.
4. Методические материалы «Тесты в образовательном процессе»: Сборник методических материалов/ Десятерик Ю.Е., Марченко М.В. – Т., 2015. – 28 с.
5. Попов Д.И., Попова Е.Д. Экспертиза качества тестовых заданий: Учебное пособие / Д.И. Попов, Е.Д. Попова. Моск. гос. ун-т печати. Москва: МГУП, 2008. 84 с.
6. Сохор А.М. Логическая структура учебного материала. Вопросы дидактического анализа. М : Педагогика, 1974. 192 с.
7. Фридман Л.М. Построение и оптимизация алгоритмов распознавания отношения принадлежности // Программированное обучение и обучающие машины. Киев, 1996. Вып. 1.
8. Аванесов В.С. Применение тестовых форм в Rasch Measurement// Педагогические измерения. URL [http:// testolog.narod.ru/Theory 49.html](http://testolog.narod.ru/Theory49.html)