

УДК 378.146

СТАНДАРТИЗАЦИЯ ТЕСТОВОГО ЗАДАНИЯ

С. А. Сафонцев, Н. Ю. Сафонцева

Южный федеральный университет (г. Ростов-на-Дону)

Поступила в редакцию 21 июля 2014 г.

Аннотация: анализируются функции тестирования в процессе формирования компетентности студента, связанные с определением валидности и надежности тестового задания; обосновывается процедура разработки тестового задания на основе проблемной ситуации с учетом результатов сравнения нормализованных статистических распределений данных репрезентации.

Ключевые слова: компетентность, тестирование, рефлексия, валидность, надежность, нормализация.

Abstract: analyzes the test function in the formation of the student's competence with determining the validity and reliability of the test task; justified procedure for development of test task based on the problem situation in the light of the comparison of normalized of statistical distributions data representation.

Key words: competence, testing, reflection, validity, reliability, normalization.

В процессе модернизации отечественной системы образования происходит болезненный переход от доминирования классической педагогики [1] к преимущественному развитию прагматических идей поведенческой психологии [2]. В современных условиях становления постиндустриальной системы образования на первый план выходит компетентность как комплексная характеристика обучающегося, которую необходимо непрерывно контролировать в процессе предоставления образовательных услуг [3]. Индикатором компетентности является рейтинг студента, учитывающий баллы, начисленные не только за правильное выполнение тестовых заданий, но и за обоснованные суждения по поводу различных способов выхода из проблемных ситуаций и самостоятельную проектную деятельность [4].

Проанализируем специфические функции теста рубежного контроля, являющегося обязательным элементом дидактического обеспечения учебного модуля. Первая из них – закрепление заинтересованности, возникающей в результате обсуждения проблемных заданий, посредством рефлексии собственных учебных достижений студента. Настойчивость при выполнении тестовых заданий и умение освободиться от затруднений эмоционального характера соответствуют конструктивной рефлексии, приводящей к успешному выполнению теста рубежного контроля. Пассивность испытуемого и неспособность преодолевать свои психологические комплексы – показатели де-

структивной рефлексии, приводящей к противоположному результату [5].

Валидность тестового задания показывает, в какой степени его выполнение способствует закреплению заинтересованности студента изучаемой дисциплиной. Сравнивая статистические распределения результатов репрезентации проблемного задания и составленного на его основе тестового задания, можно определить величину валидности последнего.

Второй функцией теста рубежного контроля является стимулирование внутренней мотивации студента. Так же, как обоснованные суждения по поводу вариантов разрешения проблемной ситуации, высказанные в процессе диалога преподавателя со студентами, выступают стимулирующим фактором возникновения заинтересованности изучаемой дисциплиной, результат тестирования уровня понимания программного материала может рассматриваться как фактор влияния на мотивацию студента. При этом простейшей формулой внутренней мотивации является «понимание в квадрате». Когда студент понял, что он понимает изученный материал, возникает желание обдумать произошедшее событие и открывающиеся перспективы, что приводит к конкретным действиям в данной области научного знания, в частности к самостоятельному выполнению проектного задания.

При условии высокой валидности тестового задания, а значит, соответствующей заинтересованности студентов изучаемой дисциплиной, степень стимулирования внутренней мотивации ха-

рактируется надежностью тестового задания, т.е. возможностью воспроизведения результатов тестирования с помощью репрезентативных выборок, составленных на основе единой процедуры стратификации. При повторном тестировании количество студентов, выбравших наиболее вероятный или единственный правильный вариант ответа, увеличивается за счет обучающихся с высокой мотивацией и уменьшается за счет тех, у кого она отсутствует. В результате медиана статистического распределения смещается незначительно.

Учитывая специфические функции теста рубежного контроля в процессе освоения учебного модуля, следует обратить внимание на две основные разновидности тестов, используемые в системе образования, – нормативные и критериально-ориентированные. Главным отличием критериального теста от нормативного является то, что в качестве интерпретационной системы отсчета используется конкретная область содержания, а не та или иная генеральная совокупность людей [6]. Именно поэтому рейтинговая система основывается на критериально-ориентированных оценочных нормах.

Поскольку рефлексия, как и функции интеллекта (внимание, память, восприятие, воображение), относится к психофизиологическим свойствам индивида, при стандартизации тестового задания необходимо опираться на нормальное статистическое распределение ошибок измерений (рис. 1).

Стандартизация тестового задания предполагает последовательное определение его валидности и надежности. Формализованным выражением валидности является степень совпадения статистических распределений результатов апробации тестового задания на представительной выборке студентов и репрезентативной экспертизы проблемного задания, которую

можно определить с помощью метода расхождения медиан [7].

Смещение медианы M_i реального распределения относительно эталона M_0 , в качестве которого рассматривается нормальное статистическое распределение, характеризует процент результатов измерений, оказавшихся между медианами:

$$\Delta M_i = (n_+ - n_-) / 2, \tag{1}$$

где ΔM_i – смещение медианы; n_+ и n_- – проценты результатов измерений (индикаторов) расположенных справа и слева от медианы M_0 (рис. 2).

Если необходимо определить расхождение медиан двух реальных распределений, то следует воспользоваться соотношением

$$R_{ij} = \Delta M_j - \Delta M_i. \tag{2}$$

Рассмотрим в качестве примера следующую проблемную ситуацию: «В процессе собеседования с претендентом на вакантную должность работодатель интересуется: 1) профессиональной квалификацией; 2) владением компетенциями; 3) компетентностью».

Проблемное задание содержит три вполне вероятных варианта разрешения ситуации неопределенности. Работодатель может заинтересоваться дипломами и сертификатами претендента на вакантную должность или предложить выпускнику университета рассказать о конкретных действиях, освоенных в процессе изучения какой-либо дисциплины и способствующих овладению определенной компетенцией (competency). Важнейшей комплексной характеристикой претендента на вакантную должность является компетентность (competence), т.е. конкурентоспособность выпускника университета, убедиться в наличии или от-

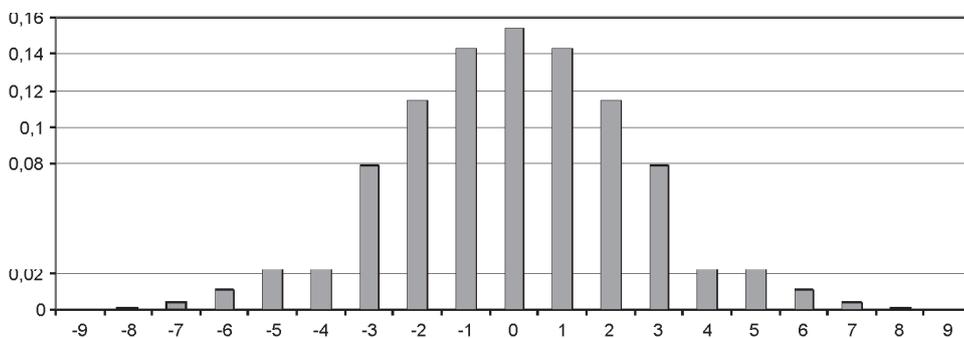


Рис. 1. Нормальное статистическое распределение

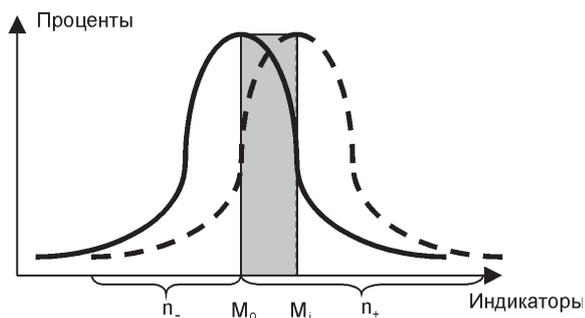


Рис. 2. Смещение медианы

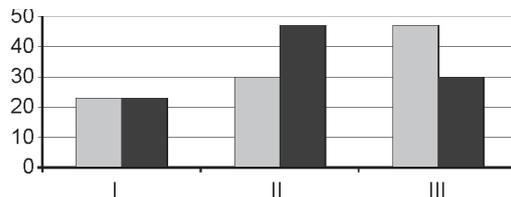
сутствии которой можно только в течение испытательного срока.

Содержание тестового задания не должно повторять проблемную ситуацию, а представлять собой конкретный пример ее реализации: «Эффективность деятельности заместителя директора предприятия проявляется в его: 1) знаниях и навыках; 2) ценностных установках; 3) психофизиологии; 4) мотивах». В предложенных вариантах ответа перечислены базовые качества компетентности, из которых в данном случае необходимо выбрать ценностные установки.

Чтобы определить расхождение медиан статистических распределений результатов репрезентации проблемного и тестового заданий, необходимо осуществить нормализацию, т.е. упорядочение вариантов ответов в соответствии с нормальным статистическим распределением. Несмотря на кажущуюся искусственность данной процедуры, она основывается на теории измерения латентных переменных [8]. Номинальная шкала, которая в нашем случае соответствует совокупности вариантов разрешения ситуации неопределенности, обычно преобразуется в ранговую шкалу, представляющую собой нарастающую последовательность порядковых номеров, которая не зависит от вероятности их выбора участниками репрезентативной выборки. Однако вполне возможно упорядочить варианты ответа таким образом, чтобы с уменьшением вероятности их выбора они располагались поочередно с обеих сторон от наиболее вероятного варианта ответа.

В результате нормализация расположения возможных вариантов ответа на проблемное задание сводится к перемещению на центральное место наиболее вероятного варианта, характеризуемого процентом отдавших ему предпочтение участников репрезентативной выборки n_3 . При этом второй по значимости n_2 вариант ответа помещается справа от центра распределения, а третий по значимости n_1 — слева. Таким образом, распределе-

ние приобретает характерный вид нормального статистического (рис. 3).

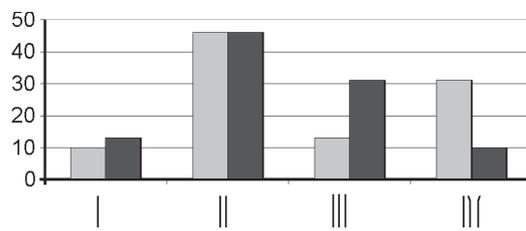


	I	II	III
■ до	23	30	47
■ после	23	47	30

Рис. 3. Нормализация результатов репрезентации проблемного задания

Если расположить эталонную медиану M_0 точно посередине второй градации, то справа и слева от нее расположатся равные половины вероятности выбора приоритетного варианта ответа n_3 . В этом случае смещение медианы статистического распределения результатов репрезентативной экспертизы определяется по формуле $\Delta M_1 = (n_2 - n_1)/2 = +3,5 \%$, записанной на основе соотношения (1). Следует обратить особое внимание на то, что порядковый номер и обозначение соответствующего процента вероятности выбора варианта ответа сохраняются после нормализации результатов репрезентации.

В процессе нормализации данных апробации тестового задания на репрезентативной выборке студентов на втором месте расположен правильный вариант ответа n_2 , наиболее популярный дистрактор n_4 переместился на третье место, а n_3 и n_1 — соответственно на первое и четвертое места (рис. 4).



	I	II	III	IV
■ до	10	46	13	31
■ после	46	13	31	10

Рис. 4. Нормализация результатов репрезентации тестового задания

Учитывая, что медиана эталонного распределения делит n_2 пополам, находим соотношение для смещения медианы реального распределения результатов апробации по формуле (1):

$$\Delta M_2 = (n_4 + n_1 - n_3) / 2 = +14,0 \%$$

Расхождение медиан нормализованных распределений результатов репрезентации проблемного (ПЗ) и тестового (ТЗ) заданий находится с помощью соотношения (2) и составляет +10,5 % (рис. 5).

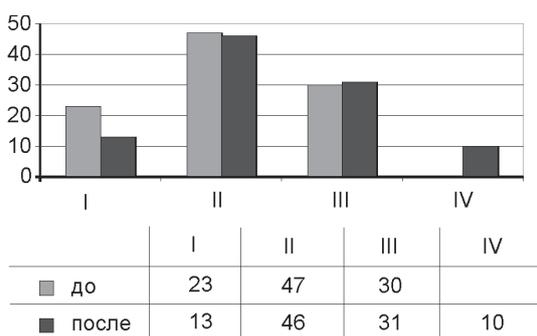


Рис. 5. Сравнение результатов репрезентации

Валидность тестового задания вычисляется как доверительная вероятность α , связанная со статистической значимостью полученного результата p соотношением:

$$\alpha = 1 - p = 1 - |R_{12}| : 100. \quad (3)$$

Подставляя в формулу (3) абсолютную величину расхождения медиан $|R_{12}| = 10,5 \%$, получаем 0,90. Это значит, что выполнение тестового задания способствует закреплению заинтересованности студента изучаемой дисциплиной с вероятностью 90 %.

Продолжая стандартизацию тестового задания, определим его надежность. Для этого потребуется по истечении некоторого времени провести повторное тестирование репрезентативной выборки студентов или сформировать новую выборку, сохранив ранее используемые стратификационные признаки (пол, место жительства, уровень доходов) и соответствующие показатели (мужской – женский, город – село, высокий – средний – низкий), образующие двенадцать страт (городские мужчины с высоким доходом, сельские женщины со средним доходом и т.д.) (рис. 6).

В данном случае нормализация распределений результатов выполнения тестового задания не требуется. Достаточно установить положение медианы эталонного распределения между второй

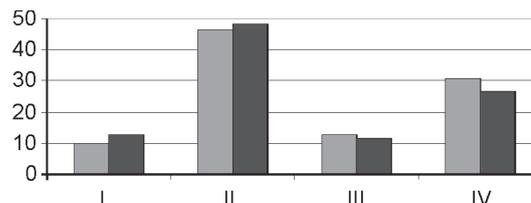


Рис. 6. Определение надежности

и третьей градациями ранговой шкалы, а затем определить расхождение медиан распределений результатов первичного и повторного тестирования. В этом случае формула смещения медианы (1) выглядит так: $\Delta M_i = (n_3 + n_4 - n_1 - n_2) / 2$. Подставляя в нее вероятности выбора различных вариантов ответа, получаем $\Delta M_1 = -6 \%$ и $\Delta M_2 = -11 \%$. В результате расхождение медиан (2): $R_{12} = \Delta M_2 - \Delta M_1 = -5 \%$. Надежность тестового задания, рассчитанная, как и валидность, с помощью соотношения (3), равна 0,95. Это значит, что в данном случае стимулирование внутренней мотивации проектной деятельности осуществляется с вероятностью 95 %.

Определение валидности и надежности тестового задания позволяет осуществить его стандартизацию, гарантирующую эффективность психолого-педагогической рефлексии собственных учебных достижений, способствующей закреплению заинтересованности изучаемой дисциплиной и стимулированию внутренней мотивации обучающегося. Каждое задание теста рубежного контроля должно пройти процедуру стандартизации с целью согласования структуры и содержания проблемных и тестовых заданий, включенных в состав учебного модуля, а также обеспечения устойчивости образовательной системы, целевой функцией которой является усвоение информации на уровне понимания [9].

В результате задания теста рубежного контроля обладают средней трудностью, каждое из них правильно выполняют 40–60 % участников репрезентативной выборки студентов университета. Если при этом их валидность примерно 0,90, а надежность не меньше 0,95, то выполнение студентом более половины заданий теста свидетельствует о конструктивной рефлексии собственных достижений и мотивации проектной деятельности.

Но это не значит, что не прошедший тестирование студент не способен самостоятельно вы-

полнить проектное задание. Просто ему придется не только следовать поставленной цели, но и воспользоваться разобранными в процессе освоения модуля средствами получения желаемого результата, рассматривая лишь другую конкретную ситуацию. Подобное выполнение проектного задания соответствует более низкому уровню компетентности, который выразится в рейтинге студента. Таким образом, использование стандартизованных тестовых заданий позволяет целенаправленно подвести студента к самостоятельному выполнению проектного задания в соответствии с его потенциальными возможностями освоения учебного модуля.

ЛИТЕРАТУРА

1. Бондаревская Е. В. Методологические проблемы становления педагогического образования университетского типа / Е. В. Бондаревская // Педагогика. – 2010. – № 9. – С. 84–88.
2. Торндайк Э. Бихевиоризм. Принципы обучения, основанные на психологии. Психология как наука о поведении / Э. Торндайк, Дж. Уотсон. – М.: АСТ-ЛТД, 1998.
3. Ефремова Н. Ф. Качество оценивания как гарантия компетентностного обучения студентов / Н. Ф. Ефремова // Высшее образование в России. – 2012. – № 11. – С. 119–125.
4. Сафонцев С. А. Принцип пакетирования рейтинговых баллов / С. А. Сафонцев, Н. Ю. Сафонцева // Образование и наука. – 2014. – № 4. – С. 146–158.
5. Atkinson R. C. Motivation and ability / R. C. Atkinson, P. M. O'Malley, W. Lens. – N. Y., 1976.
6. Glaser R. Instructional technology and the measurement of learning outcomes: some questions / R. Glaser // American Psychologist. – 1963. – № 18. – P. 519–521.
7. Сафонцев С. А. Квалиметрические измерения в процессе мониторинга структуры и содержания образовательных программ / С. А. Сафонцев, С. В. Попов, Н. Ю. Сафонцева // Стандарты и мониторинг в образовании. – 2005. – № 4. – С. 20–23.
8. Маслак А. А. Измерение латентных переменных в социально-экономических системах: теория и практика / А. А. Маслак. – Славянск-на-Кубани: Издат. центр СГПИ, 2007.
9. Сафонцев С. А. Целевая функция понятийного усвоения информации / С. А. Сафонцев, Н. Ю. Сафонцева // Alma mater (Вестник высшей школы). – 2013. – № 11. – С. 47–51.

Южный федеральный университет (Ростов-на-Дону)

Сафонцев С. А., доктор педагогических наук, профессор кафедры психологии и педагогики высшего образования

E-mail: safontsev-sa@yandex.ru

Тел.: 8(918)551-26-09

Сафонцева Н. Ю., доктор педагогических наук, профессор кафедры педагогики

E-mail: safnat67@mail.ru

Тел.: 8(918)503-91-06

South Federal University (Rostov-on-Don)

Safontsev S. A., Doctor of Pedagogical Sciences, Professor of the Psychology and Pedagogics of High Educational Department

E-mail: safontsev-sa@yandex.ru

Tel.: 8(918)551-26-09

Safontseva N. Yu., Doctor of Pedagogical Sciences, Professor of the Pedagogics Department

E-mail: safnat67@mail.ru

Tel.: 8(918)503-91-06