

УДК 004.62:519.2

М.С. МАЗОРЧУК, Е.О. БОНДАРЕНКО, В.С. ДОБРЯК

Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ»

ОЦЕНКА КАЧЕСТВА ТЕСТОВ НА ОСНОВЕ АНАЛИЗА ДИСТРАКТОРОВ ПО МЕТОДУ ПОРОГОВЫХ ГРУПП

В данной работе предлагается алгоритм оценки качества тестовых заданий на основе анализа дистракторов (вариантов ответов по тестовым заданиям) по методу пороговых групп. Данный метод предполагает исследование показателей качества теста не только для всей популяции испытуемых, но и для отдельных групп. Для каждой субпопуляции отдельно рассчитываются психометрические характеристики. Благодаря такому анализу теста можно выявить, например, некорректно сформулированные вопросы. При проведении расчетов использовались приложения статистического анализа данных: SPSS и MS Excel. Использование метода пороговых групп позволяет повысить качество педагогических тестов.

Ключевые слова: дистракторы; оценка качества тестов; пороговые группы; психометрические характеристики.

Введение

В настоящее время проблеме оценки качества педагогических тестов уделяется немало внимания. Эта тема является актуальной в связи с последними тенденциями проведения независимого тестирования для оценки уровня знаний выпускников во многих странах СНГ и Европы. В большинстве случаев определяются основные психометрические характеристики тестов, такие как сложность и дискриминация тестовых заданий, валидность и надежность теста, корреляции тестовых заданий между собой и общими результатами теста, выполняется общий анализ дистракторов¹ [1]. Для оценки этих показателей используются различные методы и модели, которые базируются на классической теории тестов и методах IRT (Item Response Theory) [2]. Однако данные методы не позволяют провести подробный качественный анализ отдельных тестовых заданий, что приводит к некорректно поставленным вопросам и содержательным ошибкам в последующих тестах. Для минимизации ошибок при разработке и экспертизе новых заданий тестов целесообразно использовать метод пороговых групп [3], который позволяет получить более подробную информацию о результатах тестирования и повысить качество

измерительных инструментов и процедур педагогических тестов. Особенно важным является проведение данного анализа для тестов, которые применяются для оценки знаний большой популяции участников – тесты Внешнего независимого оценивания (ВНО) и Тесты общей учебной компетенции (ТОУК) в Украине, Единый государственный экзамен (ЕГЭ) в России и других.

Целью данной работы является разработка алгоритма оценки качества тестовых заданий на основе анализа дистракторов по методу пороговых групп. Объектом исследования являются тесты с вопросами «закрытого» типа, т.е. с вариантами выбора правильного ответа. Предметом исследования – метод пороговых групп для анализа ответов-дистракторов.

1. Постановка задачи

Рассмотрим входные данные для анализа. Любой тест состоит из конечного числа вопросов, каждый из которых имеет несколько вариантов ответов, например А,В,С,Д, среди которых следует выбрать один или несколько правильных вариантов. Результаты ответов на тест по некоторой популяции испытуемых удобно представить в виде таблицы, где в строках отображаются варианты ответов по каждому испытуемому, а по столбцам – варианты ответов по каждому вопросу. В таблице 1 приведен пример таблицы с ответами. В таблице 2 приведен пример таблицы правильных ответов, которая используется для перекодировки ответов полученных результатов в баллы по тесту.

¹Дистракторы – варианты ответов в заданиях с выбором, не являющиеся правильными решениями, но внешне близкие к правильному решению. Использование неадекватных дистракторов искусственно ухудшает результаты тестирования, т. к. испытуемые будут введены в заблуждение. При этом в наибольшей степени пострадает наиболее «сильная» по исследуемым показателям часть испытуемых.

Таблица 1

Ответы на вопросы по тесту участников тестирования

Задания	Задание 1	Задание 2	Задание j	...	Задание 52
Номер участника						
1	A	C	...	C	...	A
2	B	B	...	C	...	D
...
i	B	A	...	C	...	B
...
n	A	D	...	B	...	A

Таблица 2

Варианты правильных ответов

Задания	Задание 1	Задание 2	Задание j	...	Задание 52
Правильные ответы	A	B	...	C	...	A

Путем перекодирования результатов, представленных в таблице 1 с помощью «ключей», которые отображены в таблице 2, формируется таблица или матрица результатов по каждому участнику (табл. 3), где a_{ij} принимает значение 1 или 0, если i -й ответ верен - присваивается 1 балл; t_i - сумма бал-

лов, набранных i -м испытуемым, где $i = \overline{1, n}$ - номер участника, а $j = \overline{1, k}$ - номер задания. Конечный результат будет выглядеть следующим образом (табл. 4).

Таблица 3

Матрица результатов

Задания	Задание 1	Задание 2	Задание j	...	Задание k	Тестовый балл i -го участника
Номер участника							
1	a_{11}	a_{12}	...	a_{1j}	...	a_{1k}	t_1
2	a_{21}	a_{22}	...	a_{2j}	...	a_{2k}	t_2
...
i	a_{i1}	a_{i2}	...	a_{ij}	...	a_{ik}	t_i
...
n	a_{n1}	a_{n2}	...	a_{nj}	...	a_{nk}	t_n

Таблица 4

Результаты ответов на вопросы по тесту

Задания Номер участника	Задание 1	Задание 2	Задание 52	Тестовый балл участника (суммарный балл по всем заданиям)
1	1	0	...	1	45
2	0	1	...	0	27
...
5245	0	0	...	0	20
...
10645	1	0	...	1	35

Идея метода пороговых групп заключается в следующем: исследование показателей качества теста проводится не только для всей популяции испытуемых, но и для отдельных групп, которые определяются установленными уровнями успешности прохождения теста. Эти уровни задаются экспертным путем исходя из предположений о количестве участников, имеющих соответствующий уровень подготовки, т.е. количестве участников, ответивших на определенное количество вопросов, которые отображают уровень подготовки испытуемого. Количество участников, как правило, задается в процентном соотношении ко всей популяции. Если рассмотреть пороговые баллы Внешнего независимого оценивания (ВНО), то можно выделить следующие «пороги»:

- 124 балла (непрофильный пороговый балл) - балл успешности прохождения теста ВНО для непрофильных предметов;
- 140 баллов (профильный пороговый балл) - балл успешности прохождения теста ВНО для профильных предметов;
- 150 баллов (средний пороговый балл) - средний балл, набранный участниками тестирования;
- 160, 175 (пороги качества 160, 175) - симметричные 140 и 124 баллам относительно среднего значения тестового балла;
- 190 (высокий пороговый балл) – показатель высокого уровня прохождения теста.

Каждый пороговый балл N определяет три пороговые группы:

- N_m - группа участников, набравшая на тестировании N баллов;

– N^+ - группа участников, набравшая на тестировании не менее N баллов;

– N^- группа участников, набравшая на тестировании меньше N баллов.

В результате применения метода пороговых групп каждый показатель качества тестовых заданий и теста в целом превращается в функцию, которая каждому пороговому баллу ставит в соответствие определенную психометрическую характеристику, например, сложность заданий $s_j = f(N_m)$ для группы участников, набравших N_m баллов, $s_j = f(N^-)$ для группы испытуемых, набравших N^- и $s_j = f(N^+)$ - для участников с N^+ .

2. Алгоритм проведения расчетов

Алгоритм расчета показателей качества теста в разрезе пороговых групп заключается в следующем.

На первом этапе, по данным таблиц 1 и 4 строится таблица сопряженности (таблица частот), где по строкам отображаются варианты ответа (дистракторы), а по столбцам – баллы t_j , набранные участниками в результате тестирования. Результаты будут иметь вид, представленный на рис. 1 (показаны фрагменты таблицы частот по одному заданию по результатам некоторого теста с баллами в шкале 100-200). На пересечении строк и столбцов отображается абсолютная частота дистракторов. Числа в последней строке и в последнем столбце (Итого) показывают суммы значений соответственно по строкам и столбцам.

		100.0	100.5	102.0	105.0	108.5	113.0	116.0	120.5	125.5	128.5	131.5	134.0	136.0	
Задание 1		30	3	1	1	3	1	1	1	3	1	5	1	2	
	A	0	16	46	84	178	306	388	612	582	596	612	543	439	
	B	0	5	10	26	58	90	119	139	142	124	95	74	45	
	C	1	25	64	165	308	559	859	1087	1186	1393	1313	1294	1223	
	D	1	30	47	115	212	388	562	735	784	872	909	818	753	
	Итого	33	79	168	391	759	1344	1929	2574	2697	2986	2934	2731	2462	

		173.0	175.0	177.0	179.0	181.0	183.5	186.0	189.0	192.0	194.5	197.0	199.0	199.5	200.0	Итого
...	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	83
	515	510	476	466	434	387	315	336	243	170	133	84	27	13	21102	
	1	1	0	0	1	0	1	0	0	0	0	0	0	0	1186	
	533	551	503	478	440	434	338	346	272	196	141	73	27	6	34718	
	510	535	495	494	424	409	370	322	263	206	149	69	40	14	26609	
	1559	1597	1474	1438	1299	1230	1024	1004	778	572	423	226	94	33	83701	

Рис. 1. Вид таблицы сопряженности для одного задания теста

Расчеты можно проводить с использованием пакета статистического анализа данных SPSS [4] или непосредственно в MS Excel [5], используя встроенный пакет анализа данных или функцию ЧАСТОТА (массив_данных; массив_интервалов).

На втором этапе, на основе таблицы сопряженности рассчитываются требуемые показатели качества теста по каждой пороговой группе. Например, сложность задания определяется по формуле:

$$s_j = \frac{c_j}{n \cdot I_j},$$

где $c_j = \sum_{i=1}^n a_{ij}$;

$I_j = \max(a_{ij})$ - максимально возможный балл за задание.

Тогда, сложность заданий для группы 124⁻ будет определяться как:

$$s_j = f(124^-) = \frac{\sum_{j=1}^{N=124^-} c_j}{124^- \cdot I_j}.$$

Пример расчета сложности заданий в MS Excel для первого варианта ответа и пороговой группы 124⁻ показан на рис. 2. Максимально возможный балл за данное задание I_j равен единице. Аналогичным образом рассчитываются сложности заданий

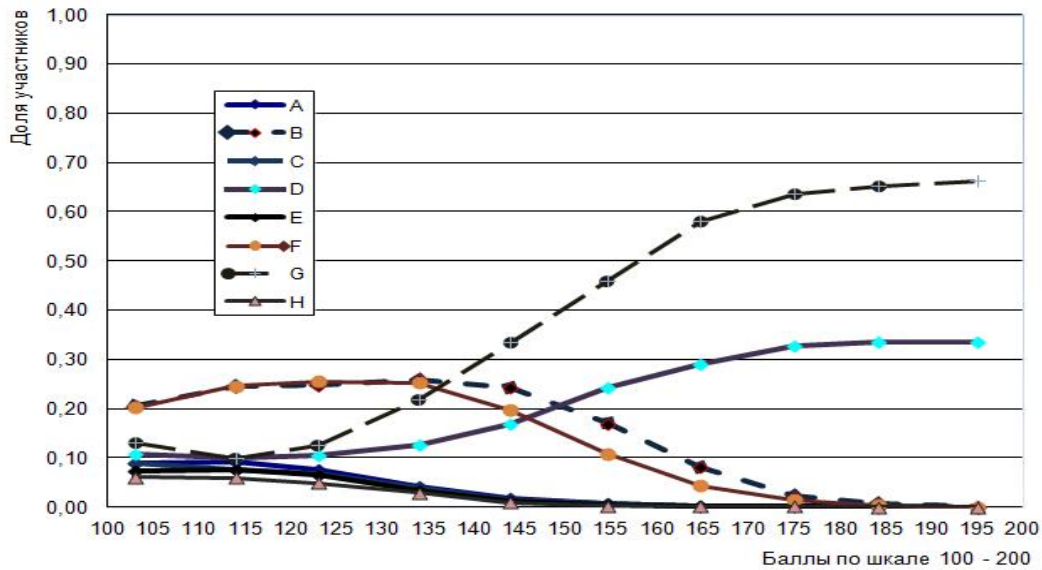
для каждой пороговой группы. По полученным результатам можно построить графики сложностей задания в зависимости от значения порога.

На рис. 3 представлены примеры серии графиков сложности вопросов по различным заданиям. Ось абсцисс - тестовые баллы по шкале 100-200; ось ординат - доля участников, получивших соответствующий суммарный тестовый балл. Каждая из кривых является графиком доли участников, выбравших соответствующий вариант ответа в зависимости от величины набранного суммарного тестового балла.

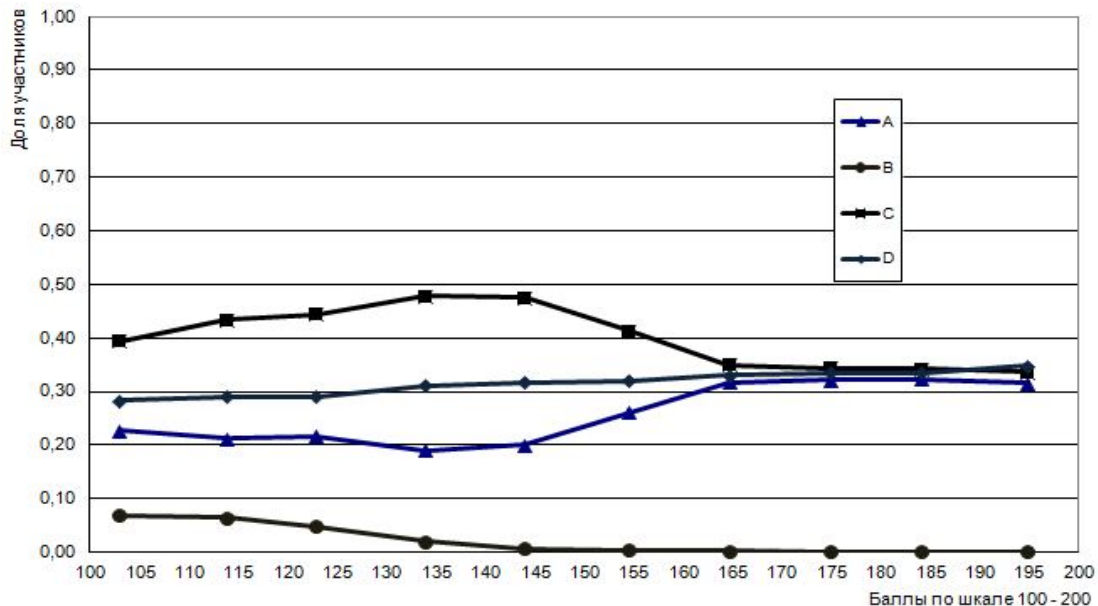
Следующим этапом является анализ полученных кривых. Как правило, для проведения качественного исследования необходима информация о содержании задания. Также привлекаются эксперты, занимающиеся разработкой тестовых заданий в соответствующей области. Однако, анализируя графики на рис. 3 и при отсутствии информации о содержании задания можно сделать выводы, что задание на рис. 3а не является сложным, так как на графике видно, что дистрактор G является правильным ответом и, начиная с группы испытуемых с баллами 140⁺, большинство выбирало данный вариант ответа. По графикам на рис. 3б можно сказать, что вопрос не был задан корректно и был достаточно сложным, так как варианты A, C и D были сильно похожи и выбирались с равной частотой, скорее всего наугад.

		100,0	100,5	102,0	105,0	108,5	113,0	116,0	120,5	125,5	128,5
Ехб		30	3	1	1	3	1	1	1	3	1
	A	0	16	46	84	178	306	388	612	582	598
	B	0	5	10	28	58	90	119	139	142	124
	C	1	25	64	165	308	559	859	1087	1186	1393
	D	1	30	47	115	212	388	562	735	784	872
	Итого	33	79	168	391	759	1344	1929	2574	2697	2988
		[100-110)	[110-120)	[120-130)	[130-140)	[140-150)	[150-160)	[160-170)	[170-180)	[180-190)	[190-200]
		103	114	123	134	144	155	165	175	184	195
	A	0,23	0,21	0,22	0,19	0,20	0,26	0,32	0,32	0,32	0,32
	B	0,07	0,06	0,05	0,02	0,01	0,00	0,00	0,00	0,00	0,00
	C	0,39	0,43	0,44	0,48	0,47	0,41	0,35	0,34	0,34	0,34
	D	0,28	0,29	0,29	0,31	0,32	0,32	0,33	0,33	0,33	0,35
		<124	>=124	>=150	>=160	>=175	>=190				
	A	=СУММ(C5:J5)/СУММ(C9:J9)	0,25	0,30	0,32	0,32	0,32				
	B		0,06	0,01	0,00	0,00	0,00				
	C		0,42	0,41	0,37	0,35	0,34				
	D		0,29	0,32	0,33	0,33	0,34				

Рис. 2. Пример расчета сложности заданий по первой пороговой группе для первого дистрактора



а



б

Рис. 3. Примеры графиков – характеристические кривые дистракторов по различным заданиям: а – график несложного задания, б – график сложного задания.

Также по результатам данного задания сложно дифференцировать участников тестирования, особенно тех, кто набрал высокие баллы.

Кроме сложности заданий, для каждой группы испытуемых можно рассчитать такие показатели качества теста, как дискриминацию тестовых заданий (D-индекс), α -Кронбаха по пороговым группам участников по всему тесту, коэффициенты корреляции правильного ответа с остальными дистракторами и др.

Заключение

Таким образом, оценка качества тестов на основе анализа дистракторов по методу пороговых групп позволяет проводить более подробный анализ тестовых заданий, оценивать содержание вопросов и корректность сформированных ответов.

Алгоритм анализа дистракторов по методу пороговых групп не является сложным, но расчеты требуют достаточно больших временных затрат, особенно при большом количестве вопросов и вари-

антов ответов. Поэтому, проведение расчетов в MS Excel не является эффективным и требуется разработка программного инструментария для проведения подобного анализа.

Литература

1. Чельщикова, М.Б. Теория и практика конструирования педагогических тестов [Текст]: учебное пособие / М.Б. Чельщикова. - М.: Логос, 2002. - 432 с.
2. Бююль, А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей [Текст] /

А. Бююль, П. Цефель. - СПб.: ДиасофтЮП, 2005. - 608 с.

3. Крокер, Л. Введение в классическую и современную теорию тестов [Текст] / Л. Крокер, Дж. Алгина. - М.: Логос, 2010. - 668 с.

4. Корнелл, П. Анализ данных в Excel. Просто как дважды два [Текст]: пер с англ. / П. Корнелл. - М.: Эксмо, 2007. - 224 с.

5. Раков, С.А. Метод пороговых групп та його використання для аналізу результатів ЗНО [Текст] / С.А. Раков // Вісник ТІМО (Тестування і моніторинг в освіті). - № 7-8. - Х.: Факт, 2013. - С. 52 - 63.

Поступила в редакцию 05.09.2013, рассмотрена на редколлегии 11.09.2013

Рецензент: д-р техн. наук, проф., зав. каф. охраны труда, стандартизации и сертификации Р.М. Трищ, Украинская инженерно-педагогическая академия, Харьков.

ОЦІНКА ЯКОСТІ ТЕСТІВ НА ОСНОВІ АНАЛІЗУ ДИСТРАКТОРІВ ПО МЕТОДУ ПОРОГОВИХ ГРУП

М.С. Мазорчук, О.О. Бондаренко, В.С. Добряк

У даній роботі пропонується алгоритм оцінки якості тестових завдань на основі аналізу дистракторів (варіантів відповідей за тестовими завданнями) за методом порогових груп. Даний метод передбачає дослідження показників якості тесту не тільки для всієї популяції випробовуваних, але і для окремих груп. Для кожної субпопуляції окремо розраховуються психометричні характеристики. Завдяки такому аналізу тесту можна виявити, наприклад, некоректно сформульовані запитання. При проведенні розрахунків використовувалися програми статистичного аналізу даних: SPSS та MS Excel. Використання методу порогових груп дозволяє підвищити якість педагогічних тестів.

Ключові слова: дистрактори; оцінка якості тестів; порогові групи; психометричні характеристики.

QUALITY ASSESSMENT TESTS BASED ON ANALYSIS DISTRACTOR ON METHOD OF CUT-OFF GROUP

M.S. Mazorchuk, O.O. Bondarenko, V.S. Dobriak

In this paper we propose an algorithm to assess the items quality of test based on analysis of distractors (answer options on the test tasks) by the threshold groups. This method involves the research indicators of test quality, not only for the whole population of subjects, but for the individual groups. For each subpopulation separately calculated psychometric characteristics. Thanks to this test analysis can reveal, for example, incorrectly worded questions. The use of threshold groups can improve the quality of educational tests. To make calculations using application of statistical data analysis: SPSS and MS Excel. Using the method of threshold groups improves quality of educational tests.

Keywords: distractors, evaluation of the quality of tests, the threshold group, psychometric characteristics.

Мазорчук Марія Сергеевна – канд. техн. наук, доцент, доцент кафедри інформатики, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков, Украина, e-mail: mazorchuk_mary@inbox.ru.

Бондаренко Елена Олеговна – аспирант кафедри інформатики, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков, Украина, e-mail: mango26.88@mail.ru.

Добряк Виктория Сергеевна - ассистент кафедри інформатики, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков, Украина, e-mail: viktoriya--13@mail.ru.