

**Анализ и совершенствование процедур тестирования выпускника
на примере программы 27.04.02 «Управление качеством»**

О.С. Шикула, А.Э. Аксенов, В.А. Кулакова

*Национальный исследовательский университет «МИЭТ»,
г. Москва, Россия,
e-mail: oshik78@mail.ru*

**Analysis and improvement of graduate testing procedures on the
example of the program 27.04.02 "Quality Management"**

O.S. Shikula, A.E. Aksenov, V.A. Kulakova

*National Research University of Electronic Technology,
Moscow, Russia,
e-mail: oshik78@mail.ru*

В настоящей статье проанализирована потенциальная польза при конструировании тестов от применения однопараметрической модели Раша для повышения объективности результатов тестирования.

Ключевые слова: компетенция, тестирование, теория педагогических измерений (IRT), латентные переменные, модель Раша.

The quality of tests, their accuracy, objectivity can be improved with the use of mathematical statistical apparatus. In this paper, we analyzed the potential benefits in constructing tests from the use of the one-parameter Rasch model to increase the objectivity of the test results.

Keywords: competence, testing, Item Response Theory (IRT), latent variables, Rasch model.

Качественное усвоение знаний тесно связано с контролем за степенью освоения материала. В последние годы стала широко применяться такая форма контроля знаний, как тестирование. В рамках реализации проекта по внедрению процедуры «рубежного контроля» в МИЭТ накопился большой объем тестовых материалов по большинству реализуемых дисциплин в вузе. Анализируя результаты проведенных тестирова-

ний, возникла идея исследовать методики составления тестов и интерпретации результатов тестирования.

В первую очередь было проанализировано, что оценивается при тестировании студентов по разным дисциплинам (курсам, модулям) учебного плана:

- уровень знаний;
- степень сформированности компетенций;
- стессоустойчивость;
- степень структурированности знаний?

Мы решили оценить степень сформированности компетенции. Это возможно осуществить через индикаторы - задания в тесте.

Возник следующий вопрос: оценить или измерить?

Оценивание подразумевает некоторую степень субъективности суждения, а хочется получить истинную информацию. Значит, надо измерять. Другой вопрос: чем измерять (инструмент и методика)?

Статистическую достоверность измеряемым переменным обеспечивает применение современной математической модели при конструировании тестов и интерпретации его результатов. Была выбрана модель Раша как наиболее применимая к образовательной деятельности.

Для того чтобы считать результаты тестирования объективной оценкой знаний студентов, необходимо быть уверенным в высоком качестве теста (его диагностической состоятельности). Это тоже обеспечивает применение модели Раша.

Датский математик Георг Раш является признанным авторитетом в современной теории педагогических измерений (IRT) [2]. В своей модели он оперирует латентными (в явном виде не измеряемыми) переменными в образовании.

Для того чтобы переменные оценивались в одной и той же шкале, он ввел интегральную шкалу логитов и ограничился двумя переменными: готовность студента и сложность тестового задания [1].

Г. Раш ввел две меры: «логит уровня знаний» Θ и «логит уровня сложности задания» δ .

$$\theta_i = \ln \frac{\text{Доля верных ответов тестируемого}}{\text{Доля неверных ответов тестируемого}},$$
$$\delta_j = \ln \frac{\text{Доля неверных ответов на } j\text{-й вопрос}}{\text{Доля верных ответов на } j\text{-й вопрос}}.$$

На практике чаще решается задача: по ответам тестируемых на задания теста оценить значения латентных параметров Θ и δ .

Модель Раша описывает вероятность успеха испытуемого как функцию одного параметра ($\Theta_i - \delta_j$), поэтому иногда ее называют однопараметрической моделью ИРТ. Эта модель позволяет строить очень интересные с точки зрения тестолога-исследователя графики. В работе [3] А.А. Маслака приведены результаты его исследований на эту тему, в том числе - графики (рис.1 - 2).

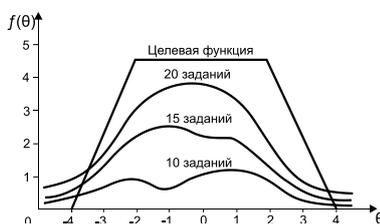


Рис.1. Информационные кривые трех тестов.
Источник: [3]

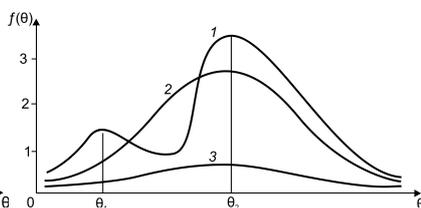


Рис.2. Информационные кривые тестов с разным количеством заданий.
Источник: [3]

Согласно А. Бирнбауму количество информации, обеспеченное j -м заданием теста в данной точке Θ_i , - это величина, обратно пропорциональная стандартной ошибке измерения данного значения Θ_i с помощью j -го задания [1]. Для описания информации, соответствующей заданию, вводится информационная функция $f(\Theta)$, приведенная на рис.3.

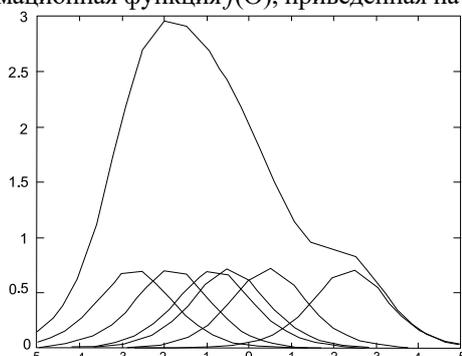


Рис.3. Информационные кривые восьми вопросов теста и всего теста для однопараметрической модели Раша.
Источник: [3]

Из информационных кривых, построенных для разных тестов (см. рис.1), для разного количества вопросов в тесте (см. рис.2) или для

каждого из заданий теста (см. рис.3), можно извлечь полезную информацию:

1. Если информационные функции для нескольких заданий накладываются, то можно оставить только одно, остальные задания исключить из теста без потери измерительных свойств всего теста.

2. Из расположения графиков видно, задания с каким уровнем сложности необходимо добавить.

3. Улучшения информационной функции всего теста можно добиться и не изменяя числа заданий в нем, для этого необходимо, например, сдвинуть задание под № 5 влево, т.е. понизить его трудность.

4. Если сравнивать информационные кривые разных тестов, то визуально можно продемонстрировать, какой из них более удачен, более информативен.

Эти возможности однопараметрической модели Раша помогут повысить объективность измерения степени сформированности компетенций в разработанном нами адаптивном тесте.

В ходе выполнения весной 2018 г. преподавательско-студенческого проекта на кафедре системной среды качества разработан адаптивный тест, ориентированный на требования профстандарта 40.062 «Специалист по качеству продукции» (в части трудовой функции G/01.7), проведена апробация разработанного адаптивного теста (участвовало 11 студентов), проведен анализ результатов тестирования и разработаны рекомендации.

Программа реализована в виде веб-приложения, что позволяет использовать ее на различных устройствах с учетом кросс-браузерности (использования различных программ веб-обозревателей). При создании использовались язык разметки HTML, языки программирования JavaScript и PHP. Программа состоит из базы данных, размещенной на MySQL сервере, и клиентской части - веб-странице. В базе данных хранится информация о дисциплинах, модулях дисциплин, вопросах и ответах. Структура и функционал программы делают ее масштабируемой и изменяемой. Для прохождения тестирования необходимо заполнить следующие поля: логин, пароль, имя, группа, тест (дисциплина) и тип тестирования (табличное или древовидное).

Если в имеющуюся версию адаптивной тестирующей программы добавить математико-статистический аппарат теории однопараметрической модели Раша, то это позволит программе стать более удобным инструментом для проектировщиков тестов и сократить время на создание качественного измерительного инструмента, повысив объективность результатов.

Библиографический список

1. *Аванесов В.С.* Item Response Theory: Основные понятия и положения // Педагогические измерения. - 2007. - № 2. - С. 3 - 28.
2. *Летова Л.В.* Исследование влияния неравномерного распределения тестовых заданий на точность измерения латентных параметров (часть 1) // Современные научные исследования и инновации. - 2014. - № 4. [Электронный ресурс]. - URL: <http://web.snauka.ru/issues/2014/04/33733> (дата обращения: 19.10.2018).
3. *Маслак А.А.* Измерение латентных переменных в социальных системах. - Славянск-на-Кубани: ИЦ КубГУ, 2012. - 432 с.