

**PSYCHOMETRIC INDICATOR
OF DOUBT
IN COMPUTER TESTS (TORFL)**

V. Serbin, Candidate of Technical Science, Project Executive of the JSC «Kaspi bank», Senior Research Associate of the National Centre of Correctional Pedagogy, Senior Lecturer of the Computer Technologies Department
Yu. Smirnova, Candidate of Education, Associate Professor
Almaty Institute of Power Engineering and Telecommunications, Kazakhstan

The substantiation of a didactic term “doubt” is given in the article. The method of the doubt level measurement (in TORFL) is presented.

Keywords: doubt, psychometry, TORFL, test

Conference participants, National championship in scientific analytics, Open European and Asian research analytics championship

**ПСИХОМЕТРИЧЕСКИЙ ПОКАЗАТЕЛЬ
СОМНЕНИЯ В КОМПЬЮТЕРНЫХ
ТЕСТАХ (TORFL)**

Сербин В.В., канд. техн. наук, управляющий проектами АО «Kaspi bank», старший научный сотрудник Национального центра коррекционной педагогики, старший преподаватель кафедры компьютерных технологий
Смирнова Ю.Г., канд. пед. наук, доцент
Алматинский университет энергетики и связи, Казахстан

В статье представлено обоснование дидактического термина «сомнение». Предложен метод измерения уровня сомнений в знаниях тестируемого (в TORFL).

Ключевые слова: сомнение, психометрия, TORFL, тест.

Участники конференции, Национального первенства по научной аналитике, Открытого Европейско-Азиатского первенства по научной аналитике

Введение. Тесты как система заданий стандартной формы, результаты которой поддаются количественному учёту,¹ получили широкое распространение в самых различных областях нашей жизни. Преимущества и недостатки использования тестов очевидны.² Ясно одно: тесты – это наиболее эффективный способ квалиметрии интеллектуальных и личностных качеств, в этом их несомненное достоинство. **Новизна** исследования состоит в разработке метода измерения уровня сомнения знаний тестируемого, а также аналитически обоснован термин «сомнение».

История термина «сомнение»: этимология, интерпретации и толкования. Сомнение – это акт мышления, колебание, раздвоение ума, ставшего перед необходимостью выбора между истинностью и ложностью какого-либо положения. В этимологическом отношении слово «сомнение» означает одновременное наличие двух или нескольких мнений об одном и том же предмете. В любом издании Толкового словаря живого великорусского языка В. Даля мы увидим, что сомнение объясняется как умение «думать надвое». А в немецком языке слово «Zweifel» (сомнение) образуется от «zwei» (два).

Первой научным направлени-

ем, специально обратившись к анализу феномена сомнения, явился античный философский скептицизм. Он вобрал в себя интеллектуальный потенциал той эпохи и высокую логическую культуру. Его стиль – это поиски, запросы и логические парадоксы. Античный скептицизм выработал особую манеру, способ противопоставлять чужие взгляды и мнения. Если для представителей других научных школ той эпохи, например, элеатов и софистов, сомнение – лишь момент познания, преодолев который они приходят к определённым положениям, то для скептиков сомнение – исходный и конечный пункт философствования, сутью которого является положение «истина не существует».³

В средние века сомнение нашло обоснование в сочинении Пьера Абеляра «Si et non» («Да и нет»). Кроме того, скептические идеи содержались в произведениях Николая Кузанского, Эразма Роттердамского, Агриппы Неттесхеймского, Мишеля Монтеня.

Дэвид Юм, английский скептик XVIII столетия, тоже представлял рассудочное мышление, развивая подходы Локка и Беркли. «Смягчённый» скептицизм Юма – это уже не скептицизм, а утверждение необходимости ограничить познание. Если Юм пытался преодолеть сомнение скептицизма, установит

пределы познания, чтобы изгнать из него противоречивость, то И. Кант стремился определить границы познания, чтобы утвердить, узаконить противоречие, доказать не отдельную его сторону, а внутреннее единство противоположностей: «Критика разума необходимо приводит в конце концов к науке; наоборот, догматическое применение разума без критики приводит к ни на чём не основанным утверждениям, которым можно противопоставить столь же ложные утверждения, стало быть, приводит к скептицизму».⁴

Фрэнсис Бэкон считал сомнение средством избавления от предрассудков и предвзятых мнений. Бэконовское сомнение противоположно сомнению скептицизма, его цель – избавить разум от призраков, оказать ему содействие в познании объективной истины.

Поворотным моментом стала интерпретация сомнения в картезианстве. По Декарту, всё можно повернуть сомнению, но несомненным является само сомнение. Сам факт сомнения не может быть опровергнут. Критерию истинности удовлетворяет несомненное сомнение: «Cogito, ergo sum». Традиционно знаменитое декартовское изречение переводят как «Я мыслю, следовательно, существую», но этот глагол можно перевести как «сомневаюсь».

1 Балыхина Т.М. Словарь терминов и понятий тестологии. – М.: Русский язык. Курсы, 2006. С.86.

2 Равен Дж. Педагогическое тестирование: проблемы, заблуждения, перспективы. – М.: Когито Центр, 1999. – 142 с.

3 Соловьёва Г.Г. О роли сомнения в познании. – Алма-Ата: Наука, 1976. – 142 с.

4 Кант И. Сочинения. <http://philosophy.ru/library/kant/01/00.html> [Дата обращения 30.05.2013].

Именно это, на наш взгляд, имел в виду Декарт, ведь именно он является основоположником методологического (в некоторых источниках – методического) сомнения, т.е. сомнения как метода установления, верифицирования истины, чем, собственно, и отличился от своих предшественников.

Сомнение – одна из безусловных форм сознания, рефлексии. Его специфика заключается и в том, что оно является логическим противоречием, единством утверждения и отрицания. Эвристичность сомнения выражается в том, что от него стремятся избавиться, т.к. неудовлетворённое сомнение мучительно, именно поэтому Декарт считал его несовершенством. Не случайно известный физик Рич. Фейман неоднократно говорил о необходимости сомнения в научном творчестве. Нет познания без вопроса, нет вопроса без сомнения.

Таким образом, в истории научной мысли выделялось скептическое сомнение и картезианское. Цель последнего – проверить на истинность все прежние положения, которые кажутся достоверными, выяснить, отражает ли противоречие в теории объективную противоречивость или оно возникло в результате логической небрежности. Именно сомнение в декартовском смысле будет рассмотрено далее.

Психометрия сомнений: математическое описание. Математическое описание сомнения как значимого компонента модели рационального выбора, в основу которой положены аксиомы независимости, транзитивности, выпуклости и монотонности индивидуальных предпочтений, некоторым образом представлено в исследовании Г.П.Виноградова и В.Н.Кузнецова.⁵ Человек строит последовательность умозаключений об адекватности своих представлений в соответствии со своим субъективным уровнем убеждённости. Он убеждён в адекватности своих представлений в ситуации

выбора типа Ω относительно цели G , если считает, что выбор способа действия C на их основе позволит её достичь. При этом он воспринимает часть характеристик X ситуации Ω ; относительно другой части он делает предположения и проявляет намерение доказать (проверить их правдоподобность); в аналогичных ситуациях выбора типа Ω , в которых он воспринимал присутствие (отсутствии) X и стремился к G , выбирая C , он всегда достигал G ; когда он наблюдал отсутствие (присутствие) X в ситуации выбора типа Ω , он никогда не выбирал C для достижения цели G на основе данного представления.

Предположение – это принимаемое по умолчанию значение наблюдаемой характеристики или описание причинно-следственной связи между наблюдаемыми характеристиками.

Представления человека характеризуются уровнем убеждённости. Уровень его убеждённости в своих представлениях о ситуации выбора типа Ω относительно цели G определяется частотой её достижения при выборе способа действия C на их основе. Оценка уровня убеждённости изменяется в пределах от нуля до единицы. Если число неудачных попыток достичь цели G при выборе способа действия C на основе представлений человека возрастает, то уровень убеждённости человека уменьшается (и наоборот), что становится для него стимулом для приложения усилий по их модификации или полной реконструкции вследствие возрастания сомнения в правдоподобности сделанных предположений. Наличие стремления проверить правильность предположений является мерой сомнений человека.

Таким образом, усилия, которые затрачивает человек для доказательства (опровержения) предположений характеризует степень его сомнений относительно представлений о ситуации выбора типа Ω при стремлении к цели G . Согласно положениям

теории психологии поведения, если уровень убеждённости, который зависит от количества подтверждений правильности выбора на основе представлений, возрастает, то стремление человека к проверке падает, так как он не видит в этом смысла. Возрастание степени сомнения является стимулом для поиска дополнительных аргументов (контраргументов).

Параметром, учитывающим эти две характеристики, является степень убеждённости, рассматриваемая как

$$S_U = U_u^\alpha \cdot (1 - S_S)^\beta \in [0, 1], \\ \alpha + \beta = 1, \quad (1)$$

где S_U – степень убеждённости; U_u – уровень убеждённости (прошлый опыт); S_S – степень сомнения человека в правильности своих представлений о ситуации выбора; α и β – коэффициенты значимости, которые человек придаёт своему опыту и необходимости поиска доказательств.⁶ Таким образом, из предыдущего выражения можно вывести формулу, выражающую степень сомнения.

Степень сомнений в современных тестах (например, TORFL, грамматический субтест на соответствие),⁷ на наш взгляд, является латентным параметром и может измеряться исключительно косвенным образом. Если степень сомнений S_S и степень уверенности U оценивать по 100-балльной шкале и выразить в %, то формула степени сомнений S_S приобретет следующий вид:

$$S_S = 100\% - U \quad (2)$$

По мнению А.В. Морозова,⁸ для измерения латентных параметров необходимы количественные характеристики. Количественными параметрами измерения степени сомнений пользователя являются следующие переменные: количество пропущенных операций; количество неподтверждённой информации; количество неоднородной информации (неоднозначной с первого раза); состояние логической цепочки (соответствие последовательности действий, уровней сложности и т.д.).

5 Г.П.Виноградов, В.Н.Кузнецов. Моделирование поведения агента с учётом субъективных представлений о ситуации выбора // Искусственный интеллект и принятие решений. № 3, 2011. С.58-72.

6 Там же, с.65-66.

7 Балыхина Т.М. Что такое русский тест? – М.: Русский язык. Курсы, 2007. С.4.

8 Морозов А.В. Психология как наука. – М., 2009. С.586.

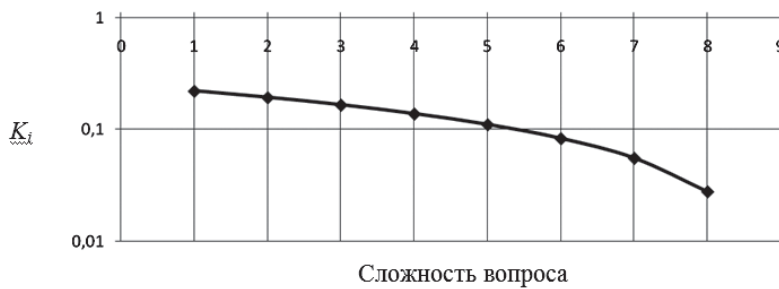


Рис. 1. График зависимости коэффициента K_i веса сомнений от текущего уровня сложности вопроса (логарифмическая).

$$V_i = \begin{cases} 0 - \text{в случае правильного ответа} \\ 1 - \text{в случае неправильного ответа} \end{cases}$$

Данная модель определяет латентный параметр, который является реакцией на принятие решения с учётом уровня сложности. Суть модели заключается в подтверждении сомнению принятого решения путём его сравнения с принятым решением подобного типа.

Определяемый латентный параметр данной модели использует следующие количественные характеристики: тематика, сложность, принятое решение.

Нами предлагается следующий метод. Организовывается тестирование из x вопросов. Все вопросы x разделены на t -группы по определённой тематике. Каждый вопрос имеет один правильный вариант ответа. В тестировании все x -вопросы поделены на n уровней сложности. Каждый вопрос имеет свой уровень сложности. Доля вопросов по уровням сложности должна удовлетворять следующему условию: для $n=3$ лёгкий – 50%, средний – 30%, сложный – 20%. Каждому уровню сложности

соответствует весовой коэффициент F (F_{max} – самый сложный, F_{max-1} – менее сложный и т.д.).

Сомнению подвергаются вопросы, на которые был дан правильный ответ всех уровней сложности, кроме первого. Сомнения при ответе на вопрос уровня F рассчитываются по ответам на вопросы той же тематики (т.е. той же группы), сложность которых ниже уровня сложности текущего вопроса.

$$S_F = \sum_{i=1}^m k_i W_i, \quad (3)$$

где m – количество уровней сложности ниже того, для которого ведётся расчёт:

$$m = F - 1 \quad (4)$$

F – весовой коэффициент текущего уровня сложности; k_i – коэффициент веса сомнений на вопрос ниже текущего уровня:



Рис.2. График зависимости коэффициента f_F суммарных сомнений от текущего уровня сложности вопроса (логарифмическая).

$$k_i = -\frac{i-m-1}{\sum_{j=1}^m j}, \quad (5)$$

W_i – суммарный вес сомнений вопроса i -го уровня,

$$W_i = \frac{1}{g} \sum_{l=1}^g V_l, \quad (6)$$

где g – количество вопросов i -го уровня сложности той же тематики, для которой ведётся расчёт сомнения; V_l – принятое решение (ответ на вопрос).

Сложность вопроса

Исходя из того, что в тесте может быть несколько вопросов одного уровня сложности, есть необходимость найти среднееквадратическое значение сомнения на каждом уровне:

$$S_i = \sqrt{\frac{\sum_{l=1}^q S_{Fi}^2}{q}} \quad (7)$$

Итоговый коэффициент, характеризующий степень сомнения, вычисляется по формуле:

$$S = \sum_{F=2}^{F_{max}} f_{F-1} \cdot S_s \quad (8)$$

где S_s – сомнение в ответах на вопрос F сложности, f_F – коэффициент веса сомнения в ответах на вопрос F сложности:

$$f_F = \frac{F}{\sum_{j=1}^m j}, \quad (9)$$

где F_{max} – максимальный вес самого высокого уровня.

Несмотря на то, что в данном методе для расчёта использовано небольшое количество характеристик (тематика, сложность, принятое решение), он даёт оценку сомнений правильно принятых решений, что, бесспорно, значимо для интерпретации конечного результата теста.⁹

Выводы. Итак, наше исследование демонстрирует оригинальный метод измерения уровня сомнений. В диагностике уровня знаний на основе компьютерного тестирования его использование позволит снизить вероятность влияния случайного правильного ответа («угадывания») на конечный результат и получить

максимально объективную оценку.

Общая схема учёта сомнений приведена в прикладном аспекте. Она была апробирована в 2008-2010 гг. в Алматинском институте энергетики и связи (ныне Алматинский университет энергетики и связи) в ходе компьютерного тестирования студентов бакалавриата по дисциплине «Русский язык» в период сессий. Использовались закрытые компьютерные тесты. Практическая значимость данного метода заключается в том, что его можно использовать для объективной оценки уровня знаний по любому предмету, где применяются тесты такого рода.

References:

1. Балыхина Т.М. Словарь терминов и понятий тестологии. – М.: Русский язык. Курсы, 2006. С.86.
2. Равен Дж. Педагогическое тестирование: проблемы, заблуждения, перспективы. – М.: Когито Центр, 1999. – 142 с.
3. Соловьёва Г.Г. О роли сомнения в познании. – Алма-Ата: Наука, 1976. – 142 с.
4. Кант И. Сочинения. <http://philosophy.ru/library/kant/01/00.html> [Дата обращения 30.05.2013].
5. Г.П.Виноградов, В.Н. Куз-

нецов. Моделирование поведения агента с учётом субъективных представлений о ситуации выбора // Искусственный интеллект и принятие решений. № 3, 2011. С.58-72.

6. Там же, с.65-66.

7. Балыхина Т.М. Что такое русский тест? – М.: Русский язык. Курсы, 2007. С.4.

8. Морозов А.В. Психология как наука. – М., 2009. С.586.

9. Serbin V.V. Methodology for Measuring the Level of Users Doubts: Start of a New Theory // Europäische Fachhochschule. European Applied Sciences. Stuttgart, № 1, 2013. P.230-233.



INTERNATIONAL SCIENTIFIC CONGRESS

***Multisectoral scientific-analytical forum
for professional scientists and practitioners***

Main goals of the IASHE scientific Congresses:

- **Promotion of development of international scientific communications and cooperation of scientists of different countries**
- **Promotion of scientific progress through the discussion comprehension and collateral overcoming of urgent problems of modern science by scientists of different countries**
- **Active distribution of the advanced ideas in various fields of science**

For additional information please contact us:
[www: http://gisap.eu](http://gisap.eu), e-mail: congress@gisap.eu

