Оценка уровня и структуры знаний испытуемых

Светлана Янченко
Уральский государственный технический университет, sy-ural@yandex.ru

Опубликовано в ж. «Педагогические Измерения №1, 2005 г.

Аннотация

В настоящей работе ставится задача получения новых оценок латентных параметров испытуемых и заданий, зависящих как от уровня подготовленности, так и от структуры знаний испытуемых. На примерах матриц тестовых результатов рассматриваются отличия предлагаемых оценок от традиционных.

Ключевые слова: педагогические измерения, тест, латентные параметры, уровень подготовленности, трудность задания, структура знаний.

1. Модель Г. Раша

В западной литературе для оценки уровня подготовленности испытуемых чаще других применяется модель Г. Раша (G. Rasch). В классическом однопараметрическом варианте этой модели предполагается, что вероятность P_{ij} верного выполнения i-м испытуемым j-го задания зависит только от присущих этому испытуемому и этому заданию параметров, характеризующих соответственно уровень подготовленности испытуемого и уровень трудности задания[1]. Эти параметры принято называть «*латентными*» (latent parameters), т.е. скрытыми, непосредственно не наблюдаемыми, но проявляющимися в процессе выполнения теста. Предполагается, что латентные параметры испытуемых и заданий являются соизмеримыми, т.е. расположенными на одной и той же шкале, что позволяет производить с ними операции сложения и вычитания.

Модель педагогического измерения Г. Раша выражается формулой:

---

[1] Rasch G. Probabilistic models for some intelligence and attainment tests. Copenhagen, Denmark, 1960. In: The Univ. Of Chicago Press, Chicago & London, 1980.

$$P_{ij} = P(\theta_i, \beta_j) = \frac{e^{\theta_i - \beta_j}}{1 + e^{\theta_i - \beta_j}}, \tag{1}$$

где θ_i – латентный параметр *i*-го испытуемого, β_j – латентный параметр *j*-го задания. Фактически в формуле (1) играет роль только один параметр – разность $\xi = \theta_i - \beta_j$, характеризующая расстояние между латентными параметрами испытуемого и задания на оси латентных переменных[2].

Единица измерения этой шкалы называется один *логит*. Величина ξ принимает значения из интервала $(-\infty; +\infty)$ и возрастает с увеличением уровня подготовленности испытуемого относительно трудности задания. При $\xi \to \infty$ вероятность верного ответа $P(\xi) = \frac{e^{\xi}}{1+e^{\xi}}$ стремится к 1 (латентный параметр испытуемого намного больше латентного параметра задания), при $\xi \to -\infty$ – к нулю. При $\xi = 0$, когда латентные параметры испытуемого и задания равны, вероятность верного выполнения задания равна 0,5. График функции $P(\xi)$, названной Г. Рашем *функцией успеха*, изображен на рис. 1.

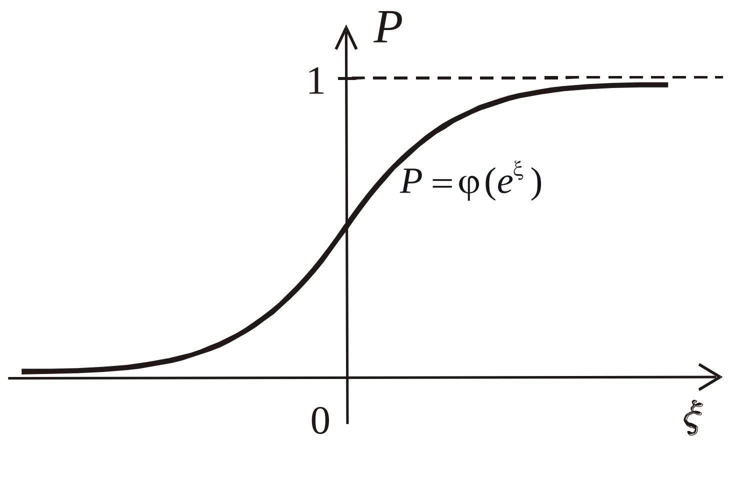

Рис. 1. График функции успеха

[2] Rasch G. Probabilistic models for some intelligence and attainment tests. Copenhagen, Denmark, 1960. In: The Univ. Of Chicago Press, Chicago & London, 1980.

2. Оценки латентных параметров: традиционный подход

В классической теории тестов в качестве первого приближения к оценкам латентных параметров рассматриваются оценки уровня подготовленности испытуемых и уровней трудности заданий, рассчитанные по матрице тестовых результатов. Пример матрицы, в которой представлены результаты тестирования небольшой группы испытуемый с небольшим же числом заданий, приводится в табл. 1[3].

В работе Luce, R.D. и Tukey J.W. такая матрица называется conjoint[4], что в переводе с английского (объединённый, совместный) означает, что матрица тестовых результатов, в общем случае, представляет результаты взаимодействия множества испытуемых, отличающихся по уровню подготовленности, и множества заданий, с различными уровнями трудности. Данная матрица представляет собой результат выполнения 13 испытуемыми 10 заданий теста.

Табл. 1. Матрица тестовых результатов

№№	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
1	1	1	1	1	0	1	1	1	1	1
2	1	1	1	1	1	1	1	1	0	0
3	1	1	1	0	1	1	1	0	1	0
4	1	1	1	1	1	1	0	0	0	0
5	0	0	0	0	1	1	1	1	1	1
6	1	1	1	0	1	0	1	0	0	0
7	1	1	0	1	1	1	0	0	0	0
8	1	1	1	1	1	0	0	0	0	0
9	1	1	1	1	0	0	0	0	0	0
10	1	0	1	1	0	0	0	1	0	0
11	1	1	0	0	0	1	0	0	0	0
12	1	0	0	0	0	0	1	0	0	0
13	0	0	1	0	0	0	0	0	0	0

[3] Аванесов В.С. Композиция тестовых заданий, С. 160. М: Центр Тестирования, 2002. - 239 с. Для данной работы понадобилось изменение профиля ответов 5-го испытуемого.
[4] Luce, R.D., Tukey, J.W. Simultaneous conjoint measurement: A new type of fundamental measurement. J. of Mathematical Psychology, 1964, 1-27.

В i-й строке матрицы записаны результаты i-го испытуемого (так называемый *профиль ответов испытуемого* i). Если он выполнил задание X_j, то на пересечении i-й строки и j-го столбца ставится 1, в противном случае – 0. Данные по всем испытуемым расположены в строках матрицы, данные по всем заданиям – в столбцах. Строки и столбцы матрицы здесь упорядочены по убыванию числа верно выполненных заданий. Таким образом, в верхних строках матрицы записаны профили ответов испытуемых с наибольшим числом верно выполненных заданий, в нижних строках – с наименьшим. В первых столбцах матрицы записаны результаты ответов на самые легкие задания теста, в последних – на самые трудные.

В статистической теории педагогических измерений уровень трудности задания определяется, исходя из доли верных либо неверных ответов, данных на задание испытуемыми. Логично за меру трудности задания принять такую оценку, которая возрастает по мере уменьшения числа испытуемых, успешно справившихся с заданием. В табл. 2 под *уровнем трудности* t_j j-го задания понимается доля неправильных ответов испытуемых на данное задание: $t_j = \frac{W_j}{N}$, где W_j – число неправильных ответов на j-е задание, N – число всех испытуемых, выполнявших тест.

Табл. 2. Матрица тестовых результатов с оценками уровней подготовленности испытуемых и трудности заданий.

№№	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	R_i	W_i	p_i	q_i	p_i/q_i	$\ln(p_i/q_i)$
1	1	1	1	1	0	1	1	1	1	1	9	1	0,9	0,1	9	2,197
2	1	1	1	1	1	1	1	1	0	0	8	2	0,8	0,2	4	1,386
3	1	1	1	0	1	1	1	0	1	0	7	3	0,7	0,3	2,333	0,847
4	1	1	1	1	1	1	0	0	0	0	6	4	0,6	0,4	1,5	0,405
5	0	0	0	0	1	1	1	1	1	1	6	4	0,6	0,4	1,5	0,405
6	1	1	1	0	1	0	1	0	0	0	5	5	0,5	0,5	1	0
7	1	1	0	1	1	1	0	0	0	0	5	5	0,5	0,5	1	0
8	1	1	1	1	1	0	0	0	0	0	5	5	0,5	0,5	1	0
9	1	1	1	1	0	0	0	0	0	0	4	6	0,4	0,6	0,667	-0,406
10	1	0	1	1	0	0	0	1	0	0	4	6	0,4	0,6	0,667	-0,406
11	1	1	0	0	0	1	0	0	0	0	3	7	0,3	0,7	0,429	-0,847

12	1	0	0	0	0	0	1	0	0	0	2	8	0,2	0,8	0,25	-1,386
13	0	0	1	0	0	0	0	0	0	0	1	9	0,1	0,9	0,111	-2,198
R_j	11	9	9	7	7	7	6	4	3	2	65					
W_j	2	4	4	6	6	6	7	9	10	11						
s_j	0,846	0,692	0,692	0,538	0,538	0,538	0,462	0,308	0,231	0,154						
t_j	0,154	0,308	0,308	0,462	0,462	0,462	0,538	0,692	0,769	0,846						
t_j/s_j	0,182	0,444	0,444	0,857	0,857	0,857	1,167	2,25	3,333	5,5						
$\ln(t_j/s_j)$	-1,705	-0,811	-0,811	-0,154	-0,154	-0,154	0,154	0,811	1,204	1,705						

Аналогично можно ввести оценку уровня легкости задания $s_j = \dfrac{R_j}{N}$ как долю правильных ответов, где R_j – число правильных ответов испытуемых на j-е задание теста. Поскольку $W_j + R_j = N$, то

$$t_j + s_j = 1, \qquad (2)$$

т.е. в статистическом подходе оценки трудности и легкости заданий жестко связаны между собой линейным соотношением (2) (см. табл. 2).

Уровень подготовленности i-го испытуемого p_i определяется из той же матрицы по формуле $p_i = \dfrac{R_i}{k}$, где R_i – число правильных ответов этого испытуемого на задания теста, k – число заданий в тесте. Уровень подготовленности есть не что иное, как относительная частота верных ответов данного испытуемого на задания теста. Симметричной оценкой является уровень неподготовленности q_i как относительная частота, или доля, неверных ответов: $q_i = \dfrac{W_i}{k}$ где W_i – число неверных ответов этого испытуемого (см. табл. 2). Справедливо соотношение

$$p_i + q_i = 1, \qquad (3)$$

т.е. оценки p_i и q_i не являются независимыми друг от друга.

Г. Раш[5] в качестве показателя подготовленности испытуемого брал отношение $p_i : q_i$, рассматривая его как *шансы* i-го испытуемого верно

[5] Rasch G. Probabilistic models for some intelligence and attainment tests. Copenhagen, Denmark, 1960. In: The Univ. Of Chicago Press, Chicago & London, 1980.

выполнить задание: если относительные частоты соответственно верного и неверного выполнения заданий равны 0,9 и 0,1, то шансы на успех равны $0{,}9:0{,}1$ или $9:1$. Отношение $p_i:q_i$ иногда называют *потенциалом* испытуемого.

В модели измерения Г. Раша уровень подготовленности испытуемого и уровень трудности задания рассматриваются как латентные параметры, измеряемые в логитах. В качестве некоторой оценки латентного параметра θ_i подготовленности i-го испытуемого, или первого приближения к значению латентного параметра, рассматривается натуральный логарифм отношения $p_i:q_i$, или потенциала. В табл. 2 эта оценка обозначена $\tilde{\theta}_i$:

$$\tilde{\theta}_i = \ln(p_i/q_i).$$

Через $\tilde{\beta}_j$ обозначена оценка латентного параметра задания, вычисляемая как логарифм потенциала j-го задания:

$$\tilde{\beta}_j = \ln(p_j/q_j).$$

Для рассматриваемой матрицы тестовых результатов оценки латентных параметров заданий распределены на отрезке $[-1{,}71; 1{,}71]$, а оценки латентных параметров испытуемых – на отрезке $[-2{,}20; 2{,}20]$ (см. табл. 2). Данные оценки обладают следующими свойствами:

1. Оценки $\tilde{\theta}_i$ и $\tilde{\beta}_j$ однозначно зависят от числа верно выполненных заданий R_i и R_j, т.е. от *исходного балла* испытуемых и заданий.
2. Оценки $\tilde{\theta}_i$ и $\tilde{\beta}_j$ монотонно зависят от исходного балла испытуемых и заданий (при уменьшении p_i увеличивается q_i, что ведет к уменьшению потенциала $p_i:q_i$), вследствие чего сохраняется порядок расположения как испытуемых по уровню подготовленности, так и заданий по уровню трудности.

3. Оценки $\tilde{\theta}_i$ испытуемых не зависят от качества выполненных ими заданий (были ли эти задания трудными или легкими), а только от их количества.
4. Оценки $\tilde{\beta}_j$ заданий не зависят от уровня подготовки справившихся с ними испытуемых (были ли они хорошо или плохо подготовлены), а только от их количества.

Иными словами, оценка $\tilde{\theta}_i$ латентного параметра испытуемого не учитывает *структуру* его *знаний.* Структура знаний зависит от порядка расположения единиц и нулей в профиле ответов испытуемого.

Рассмотрим следующий пример: двое испытуемых выполняют одни и те же задания теста, но первый выполнил верно только 6 самых легких заданий, а второй – только 6 самых трудных. В матрице тестовых результатов со столбцами, упорядоченными по возрастанию статистической трудности заданий, их профили ответов выглядят, соответственно,

1 1 1 1 1 1 0 0 0 0

и

0 0 0 0 1 1 1 1 1 1.

В этом случае можно сказать, что первый испытуемый обладает правильной структурой знаний: он выполняет легкие задания теста и не справляется с трудными. Второй испытуемый верно выполнил трудные задания, но не смог выполнить легкие. Причины такого результата могут быть разными. И всё-таки непонятно, как при такой структуре знаний оценивать действительный уровень подготовленности? Исходный тестовый балл обоих испытуемых равен 6, и их оценки θ_i одинаковы. Но профили ответов разные. Чьи знания полнее, кому и почему следует отдать предпочтение? Такая проблема возникает и при попытках шкалирования результатов ЕГЭ, что вызывает множество нареканий со стороны детей и родителей.

Приведенный пример представляет собой редкий случай, но он наводит на мысль о необходимости учитывать не только уровень подготовленности, но и структуру знаний испытуемого. Это можно сделать, если принять к рассмотрению показатель трудности выполненных испытуемым заданий.

3. Постановка первой задачи

Пусть дана матрица результатов тестирования $A=(a_{ij})$ размера $N\times k$, где N – число испытуемых, k – число заданий в тесте,

$i=1,\ldots, N;\ j=1,\ldots,k.$

Примером такой матрицы A является матрица из табл.1, если убрать первую строку с номерами заданий и первый столбец с номерами испытуемых (рис. 2). Данная матрица представляет собой результат взаимодействия 13 испытуемых с 10 заданиями теста. В этой матрице в фокусе рассмотрения находятся испытуемые, поэтому элемент матрицы равен 1, если испытуемый успешно справился с заданием, и 0, если не справился.

Наряду с матрицей A можно рассмотреть матрицу $\bar{A}$, представляющую собой тот же результат взаимодействия испытуемых с заданиями (см. рис. 2). В фокусе рассмотрения матрицы $\bar{A}$ находятся задания, а не испытуемые, поэтому элемент $\bar{a}_{ij}$ матрицы $\bar{A}=(\bar{a}_{ij})$ равен 1, если i-й испытуемый не справился с j-м заданием (задание оказалось «сильнее»), и 0, если испытуемый выполнил задание. Элементы матриц A и $\bar{A}$ связаны соотношением: $\bar{a}_{ij}=1-a_{ij}$.

Исходный балл R_i i-го испытуемого равен сумме элементов i-й строки матрицы A, а исходный балл W_j j-го задания – сумме элементов j-го столбца матрицы $\bar{A}$:

$$R_i = a_{i1} + a_{i2} + \ldots + a_{ik}\,, \qquad W_j = \bar{a}_{1j} + \bar{a}_{2j} + \ldots + \bar{a}_{Nj}\,. \tag{4}$$

Рассмотренные выше оценки уровня подготовленности p_i и уровня трудности задания t_j пропорциональны этим исходным баллам R_i и W_j.

Цель дальнейшего рассмотрения – найти такие аналоги оценок p_i и t_j, в которых учитываются соответственно уровни трудности заданий и подготовленности испытуемых. Подобные оценки могут быть получены небольшой корректировкой формул (4), а именно, использованием вместо исходных – взвешенных баллов.

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \bar{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Рис. 2. Пример матриц A и $\bar{A}$

Для каждого испытуемого с номером i, $i = 1, \ldots, N$, ищется оценка его уровня подготовленности p^*_i, лежащая на отрезке от 0 до 1; эта оценка будет аналогом рассмотренной выше оценки p_i – доли правильных ответов испытуемого на задания теста. При этом для испытуемого, выполнившего все задания теста, оценка p^*_i должна быть равна 1, а для испытуемого, не выполнившего ни одного задания – нулю. Для каждого задания теста полезно также найти оценку трудности t^*_j от 0 до 1, причем для заданий, которые не выполнил никто, $t^*_j = 1$, а для слишком простых, с которыми все справились, $t^*_j = 0$.[6]

Предположим, что существуют истинные, «идеальные» оценки трудностей заданий теста $t^*_1, t^*_2, \ldots, t^*_k$. Пусть тогда оценка уровня подготовленности i-го испытуемого p^*_i будет пропорциональна сумме оценок трудностей тех заданий теста, которые он выполнил:

$$p^*_i = \alpha \cdot (a_{i1} \cdot t^*_1 + a_{i2} \cdot t^*_2 + \ldots + a_{ik} \cdot t^*_k); \tag{5}$$

здесь числа $a_{i1}, a_{i2}, \ldots, a_{ik}$ составляют строку ответов (из 0 и 1) в матрице ответов A, т.е. профиль ответов i-го испытуемого. Коэффициент пропорциональности α выберем таким, чтобы $p^*_i \in [0,1]$: $\alpha = \dfrac{1}{t^*_1 + t^*_2 + \ldots + t^*_k}$.

С другой стороны, вычисленные по формуле (5) оценки уровней подготовки испытуемых – «правильные» оценки, учитывающие трудность выполненных заданий; нельзя ли теперь потребовать, чтобы истинная оценка трудности j-го задания t^*_j была пропорциональна сумме оценок уровней подготовленности испытуемых, которые оказались «слабее» его – не справились с ним?

$$t^*_j = \beta \cdot \left((1-a_{1j}) \cdot p^*_1 + (1-a_{2j}) \cdot p^*_2 + \ldots + (1-a_{Nj}) \cdot p^*_N\right), \tag{6}$$

или $t^*_j = \beta \cdot (\bar{a}_{1j} \cdot p^*_1 + \bar{a}_{2j} \cdot p^*_2 + \ldots + \bar{a}_{Nj} \cdot p^*_N)$,

[6] В действительности таких экстремальных заданий может и не быть, речь идет о предельном случае.

где $\beta = \dfrac{1}{p^*_1 + p^*_2 + \ldots + p^*_N}$.

На языке математики можно написать, что ищется вектор-столбец уровней подготовленности испытуемых $\mathbf{p}^* = (p^*_1, p^*_2, \ldots, p^*_N)^T$ и вектор-столбец уровней трудности заданий $\mathbf{t}^* = (t^*_1, t^*_2, \ldots, t^*_k)^T$ с действительными неотрицательными координатами, но не равные тождественно нулю.[7]

Обозначим через $\|\mathbf{p}^*\|$ и $\|\mathbf{t}^*\|$ нормы этих векторов, равные суммам координат: $\|\mathbf{p}^*\| = p^*_1 + p^*_2 + \ldots + p^*_N$, $\|\mathbf{t}^*\| = t^*_1 + t^*_2 + \ldots + t^*_k$.

Используя операцию умножения матрицы на вектор и введенную выше матрицу $\overline{A}$, запишем формулы (5) и (6) вместе с условиями на векторы $\mathbf{p}^*$ и $\mathbf{t}^*$ в виде

$$\mathbf{p}^* = \frac{1}{\|\mathbf{t}^*\|} A\mathbf{t}^*, \quad \mathbf{t}^* = \frac{1}{\|\mathbf{p}^*\|} \overline{A}^T \mathbf{p}^*, \quad \mathbf{p}^* \geq 0, \quad \mathbf{t}^* \geq 0, \quad \mathbf{p}^* \neq 0, \quad \mathbf{t}^* \neq 0. \qquad (7)$$

Назовем задачу нахождения векторов $\mathbf{p}^*$ и $\mathbf{t}^*$, удовлетворяющих условию (7), ***задачей об определении способности испытуемых и трудности заданий***, или, коротко, задачей об определении способности.

Координаты искомого вектора $\mathbf{p}^*$ можно назвать *уровнями способности* испытуемых. Способность i-го испытуемого p^*_i характеризует уровень его конкурентной подготовленности, т.е. его способность выполнять задания, вызывающие затруднения у других испытуемых[8]. Это оценка подготовленности испытуемого, отличная от традиционной доли верных ответов на задания теста.

Координаты искомого вектора $\mathbf{t}^*$ отражают меру трудности заданий, отличную от введенной ранее доли неверных ответов испытуемых. Оценка j-го задания t^*_j характеризует уровень его конкурентной невыполнимости,

---

[7] $\ldots^T$ обозначает операцию транспонирования.

[8] Подобная трактовка оценок p^*_i и ниже t^*_j следует из чисто математических соображений, которые здесь не приводятся в целях экономии места.

т.е. уровень его невыполнимости при выполнимости других заданий. Эта оценка также может служить мерилом уровня трудности задания.

4. Постановка второй задачи

Если матрица A показывает успех испытуемых, то матрица $\bar{A}$ – «успех» заданий, что означает неуспешность испытуемых (см. рис. 2). Поэтому наряду с рассмотренной выше задачей об определении способности испытуемых и трудности заданий, записанной для матрицы A, логично поставить вопрос о симметричной задаче нахождения противоположных характеристик испытуемых и заданий, записанной для матрицы $\bar{A}$. Цель второй задачи – найти оценки *неподготовленности* испытуемых, аналогичные традиционным оценкам $q_i = 1 - p_i$ (относительная частота неверных ответов), и оценки *легкости* заданий, аналогичные оценкам $s_j = 1 - t_j$ (относительная частота выполнения задания испытуемыми).

Ставится задача о нахождении векторов-столбцов $\mathbf{q}^* = (q_1^*, q_2^*, \ldots, q_N^*)^T$ и $\mathbf{s}^* = (s_1^*, s_2^*, \ldots, s_k^*)^T$, действительных и не равных тождественно нулю, отражающих соответственно уровни *неспособности* испытуемых и *легкости* заданий.

Оценка q_i^* неспособности i-го испытуемого пропорциональна сумме оценок слабости тех заданий, с которыми он не сумел справиться:

$$q^*_i = \bar{\alpha} \cdot \left(\bar{a}_{1j} \cdot s^*_1 + \bar{a}_{2j} \cdot s^*_2 + \ldots + \bar{a}_{kj} \cdot s^*_k \right), \tag{8}$$

где $\bar{\alpha} = \dfrac{1}{s^*_1 + s^*_2 + \ldots + s^*_k}$. Оценка t^*_j легкости j-го задания пропорциональна сумме оценок неспособности испытуемых, которые справились с этим заданием:

$$s^*_j = \bar{\beta} \cdot \left(a_{1j} \cdot q^*_1 + a_{2j} \cdot q^*_2 + \ldots + a_{Nj} \cdot q^*_N \right), \tag{9}$$

где $\bar{\beta} = \dfrac{1}{q^*_1 + q^*_2 + \ldots + q^*_N}$.

Коэффициенты пропорциональности $\bar{\alpha}$ и $\bar{\beta}$ выбираются из тех соображений, чтобы и q_i^*, и t^*_j принадлежали отрезку $[0,1]$. При этом для испытуемого, выполнившего все задания теста, оценка слабости q^*_i равняется 0, а для испытуемого, не выполнившего ни одного задания, равняется 1. Для трудных заданий, которые не выполнил никто, оценка слабости s^*_j равна 0, а для слишком простых, с которыми все справились, $s^*_j = 1$.

Математически задача (8), (9) записывается так:

$$\mathbf{q}^* = \frac{1}{\|\mathbf{s}^*\|}\bar{A}\mathbf{s}^*, \quad \mathbf{s}^* = \frac{1}{\|\mathbf{s}^*\|}A^T\mathbf{s}^*, \quad \mathbf{q}^* \geq 0, \quad \mathbf{s}^* \geq 0, \quad \mathbf{q}^* \neq 0, \quad \mathbf{s}^* \neq 0. \quad (10)$$

Назовем задачу нахождения векторов $\mathbf{q}^*$ и $\mathbf{s}^*$, удовлетворяющих условию (10), ***задачей об определении неспособности испытуемых и легкости заданий***, или, коротко, задачей об определении неспособности.

Неспособность испытуемого отражает уровень его конкурентной неподготовленности, т.е. неспособность выполнять задания, с которыми успешно справляются другие. Слабость задания есть характеристика его конкурентной выполнимости, т.е. выполнимости при невыполнимости других заданий.

5. Существование и единственность решения

Исследование задач об определении способности (7) и об определении неспособности (10) показало, что решение обеих поставленных задач существует и единственно. Теоретические условия существования и единственности очень необременительны.[9] В

[9] Янченко С.И. Оценка уровня подготовленности: непараметрический подход // Вопросы тестирования в образовании. 2002. № 3. С. 38-50.

практической ситуации массового тестирования, как показал анализ представительной выборки из базы данных централизованного тестирования 2001 и 2002 гг., предоставленной Центром тестирования МО РФ, решение существует и единственно всегда.[10]

6. Решение задач и его особенности

В данной работе в целях экономии места вычислительные вопросы нахождения решения поставленных задач (7) и (10) опускаются. Искомые решения, т.е. координаты векторов $\mathbf{p}^*$, $\mathbf{q}^*$, $\mathbf{t}^*$ и $\mathbf{s}^*$, приведены в табл. 3. Эти же оценки вместе с традиционными долями верных и неверных ответов приведены в табл. 4, для сравнения. Наименованиями со звездочкой («легкость*», «потенциал*») в таблицах обозначены оценки, полученные в рамках решения задач (7) и (10), в противопоставлении традиционным оценкам, полученным на основе долей верно и неверно выполненных заданий теста. Словами «оценка ЛП» и «оценка ЛП*» обозначены оценки латентных параметров в рамках модели Раша, полученные, соответственно, в рамках традиционного и предлагаемого подходов.

Табл. 3. Матрица тестовых результатов с оценками уровней подготовленности испытуемых и трудности заданий, предлагаемый подход

№№	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	R_i	W_i	p^*_i	q^*_i	p^*_i/q^*_i	$\ln(p^*_i/q^*_i)$
1	1	1	1	1	0	1	1	1	1	1	9	1	0,908	0,102	8,902	2,186
2	1	1	1	1	1	1	1	1	0	0	8	2	0,665	0,074	8,986	2,196
3	1	1	1	0	1	1	1	0	1	0	7	3	0,585	0,192	3,047	1,114
4	1	1	1	1	1	1	0	0	0	0	6	4	0,434	0,23	1,887	0,635
5	0	0	0	0	1	1	1	1	1	1	6	4	0,732	0,565	1,296	0,259
6	1	1	1	0	1	0	1	0	0	0	5	5	0,358	0,337	1,062	0,060
7	1	1	0	1	1	1	0	0	0	0	5	5	0,366	0,378	0,968	-0,032
8	1	1	1	1	1	0	0	0	0	0	5	5	0,36	0,333	1,081	0,078
9	1	1	1	1	0	0	0	0	0	0	4	6	0,268	0,435	0,616	-0,484

[10] Янченко С.И., Панов В.Г. Основные результаты НИР «Сравнительное исследование условий применимости в РФ методик тестового измерения, основанных на согласованных характеристиках силы и слабости и на модели Раша» // Итог. конф. по программе «Педагогические измерения в России» 27-29 сентября, 2002 г.: Тез. докл. М., 2002 г. С. 69-75.

10	1	0	1	1	0	0	0	1	0	0	4	6	0,34	0,504	0,675	-0,394
11	1	1	0	0	0	1	0	0	0	0	3	7	0,175	0,58	0,302	-1,198
12	1	0	0	0	0	0	1	0	0	0	2	8	0,136	0,716	0,190	-1,661
13	0	0	1	0	0	0	0	0	0	0	1	9	0,068	0,852	0,080	-2,528
R_j	11	9	9	7	7	7	6	4	3	2	65	Исходный балл исп.	**Способность**	**Неспособность**	**Потенциал***	**Оценка ЛП***
W_j	2	4	4	6	6	6	7	9	10	11						
s^*_j	0,733	0,502	0,577	0,388	0,398	0,4	0,375	0,235	0,162	0,126						
t^*_j	0,148	0,236	0,261	0,381	0,351	0,284	0,373	0,51	0,587	0,696						
t^*_j/s^*_j	0,202	0,470	0,452	0,982	0,882	0,71	0,995	2,170	3,623	5,524						
$\ln(t^*_j/s^*_j)$	-1,60	-0,755	-0,793	-0,018	-0,126	-0,342	-0,005	0,775	1,287	1,709						

См. далее табл. 4

Табл. 4. Матрица тестовых результатов с оценками уровней подготовленности испытуемых и трудности заданий.
Сравнение традиционного и предлагаемого подходов

	№№	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	R_i	W_i	p_i	p^*_i	q_i	q^*_i	p_i/q_i	p^*_i/q^*_i	$\ln(p_i/q_i)$	$\ln(p^*_i/q^*_i)$
	1	1	1	1	1	0	1	1	1	1	1	9	1	0,9	0,908	0,1	0,102	9	8,902	2,197	2,186
	2	1	1	1	1	1	1	1	1	0	0	8	2	0,8	0,665	0,2	0,074	4	8,986	1,386	2,196
	3	1	1	1	0	1	1	1	0	1	0	7	3	0,7	0,585	0,3	0,192	2,333	3,047	0,847	1,114
	4	1	1	1	1	1	1	0	0	0	0	6	4	0,6	0,434	0,4	0,23	1,5	1,887	0,405	0,635
	5	0	0	0	0	1	1	1	1	1	1	6	4	0,6	0,732	0,4	0,565	1,5	1,296	0,405	0,259
	6	1	1	1	0	1	0	1	0	0	0	5	5	0,5	0,358	0,5	0,337	1	1,062	0	0,060
	7	1	1	0	1	1	1	0	0	0	0	5	5	0,5	0,366	0,5	0,378	1	0,968	0	-0,032
	8	1	1	1	1	1	0	0	0	0	0	5	5	0,5	0,36	0,5	0,333	1	1,081	0	0,078
	9	1	1	1	1	0	0	0	0	0	0	4	6	0,4	0,268	0,6	0,435	0,667	0,616	-0,406	-0,484
	10	1	0	1	1	0	0	0	1	0	0	4	6	0,4	0,34	0,6	0,504	0,667	0,675	-0,406	-0,394
	11	1	1	0	0	0	1	0	0	0	0	3	7	0,3	0,175	0,7	0,58	0,429	0,302	-0,847	-1,198
	12	1	0	0	0	0	0	1	0	0	0	2	8	0,2	0,136	0,8	0,716	0,25	0,190	-1,386	-1,661
	13	0	0	1	0	0	0	0	0	0	0	1	9	0,1	0,068	0,9	0,852	0,111	0,080	-2,198	-2,528
												Исходный балл испытуемого		Доля верных ответов	**Способность**	Доля неверных ответов	**Неспособность**	Потенциал	**Потенциал***	Оценка ЛП	**Оценка ЛП***
	R_j	11	9	9	7	7	7	6	4	3	2										
	W_j	2	4	4	6	6	6	7	9	10	11										
Доля верных ответов	s_j	0,846	0,692	0,692	0,538	0,538	0,538	0,462	0,308	0,231	0,154										
Легкость*	s^*_j	0,733	0,502	0,577	0,388	0,398	0,4	0,375	0,235	0,162	0,126										
Доля неверн. ответов	t_j	0,154	0,308	0,308	0,462	0,462	0,462	0,538	0,692	0,769	0,846										
Трудность*	t^*_j	0,148	0,236	0,261	0,381	0,351	0,284	0,373	0,51	0,587	0,696										
Потенциал	t_j/s_j	0,182	0,444	0,444	0,857	0,857	0,857	1,167	2,25	3,333	5,5										
Потенциал*	t^*_j/s^*_j	0,202	0,470	0,452	0,982	0,882	0,71	0,995	2,170	3,623	5,524										
Оценка ЛП	$\ln(t_j/s_j)$	-1,705	-0,811	-0,811	-0,154	-0,154	-0,154	0,154	0,811	1,204	1,705										
Оценка* ЛП	$\ln(t^*_j/s^*_j)$	-1,60	-0,755	-0,793	-0,018	-0,126	-0,342	-0,005	0,775	1,287	1,709										

6.1. Оценки способности и неспособности испытуемых

Однозначная зависимость оценки способности испытуемого от набранного им исходного тестового балла отсутствует. Как правило, по мере увеличения исходного балла испытуемого оценка его способности возрастает (см. табл. 3, 4). Исключением явилась оценка способности испытуемого 5, верно выполнившего 6 последних, самых трудных, заданий теста. Эта оценка оказалась больше, чем оценки способности 3-го и 2-го испытуемых, выполнивших соответственно 7 и 8 заданий теста. Причина высокой оценки способности 5-го испытуемого кроется в общей суммарной трудности выполненных им заданий, а трудность этих заданий оценивалась мерой способности испытуемых, не справившихся с ними (см. п. 3). Таким образом, в оценке способности 5-го испытуемого оказалась учтена его способность успешно выполнять задания, с которыми не справляются другие испытуемые.

Та же закономерность прослеживается и в оценках способности испытуемых, набравших одинаковые исходные баллы (4-го и 5-го; 6-го, 7-го и 8-го; 9-го и 10-го). Во всех случаях оценка способности тем больше, чем более трудные задания выполнил испытуемый.

Структура знаний в оценке способности не учитывается; наоборот, чем более неправильную структуру знаний демонстрирует испытуемый, тем более высокую оценку способности он получает. Поэтому при определении уровня подготовленности обязательно следует учитывать и другую сторону медали – оценку, выставляемую по невыполненным заданиям теста, т.е. оценку неспособности испытуемого.

Оценки неспособности испытуемых независимы от оценок их способности. В отличие от традиционного подхода, когда суммы долей верных и неверных ответов равны 1 (см. формулу (3)), сумма оценок способности и неспособности испытуемого может быть как больше, так и меньше 1 (см. табл. 3, 4).

Общей тенденцией является возрастание оценок неспособности по мере увеличения числа невыполненных заданий теста. Порядок нарушается для испытуемых, не справившихся с заданиями, которые должны были оказаться им по силам. Так, оценка неспособности 1-го испытуемого, успешно справившегося с 9 заданиями теста, но не выполнившего 5-е задание, достаточно легкое, судя по результатам других испытуемых, оказалась больше, чем оценка неспособности 2-го испытуемого, выполнившего 8 заданий. Оценка неспособности нашего «неправильного» 5-го испытуемого, выполнившего 6 заданий, больше, чем оценки неспособности 6-го, 7-го, 8-го, 9-го и 10-го испытуемых, справившихся всего с 5 или 4 заданиями. Это объясняется тем, что среди невыполненных 5-м испытуемым заданий оказались исключительно легкие, с которыми без труда справились остальные испытуемые.

Таким образом, оценка неспособности тем больше, чем более легкие задания не сумел выполнить испытуемый; в данном случае легкость оценивается по результатам выполнения заданий всеми испытуемыми. При одинаковом исходном балле оценка неспособности учитывает структуру знаний испытуемого: чем более правильную структуру знаний он демонстрирует, тем меньше окажется его оценка неспособности.

6.2. Оценки трудности и легкости заданий

Оценка трудности заданий в рамках предлагаемого подхода зависит от количества и «качества», т.е. уровня подготовленности, испытуемых, которые не сумели с ним справиться. Для рассматриваемого примера оценка трудности задания возрастает с увеличением его исходного балла, т.е. количества испытуемых, не выполнивших это задание (см. табл. 3. 4). Данная закономерность нарушается лишь для заданий X_4 и X_7, что связано с большей суммарной оценкой способности испытуемых, не выполнивших задание X_7, по сравнению с той же величиной для задания X_4.

Напротив, оценка легкости задания при возрастании его исходного балла обычно уменьшается, что и наблюдается для данной матрицы тестовых результатов. Оценка легкости задания получается как сумма оценок неспособности испытуемых, которые смогли с ним справиться.

Сравнивая 2 задания с одинаковым исходным баллом, X_2 и X_3, и рассматривая только различия в их выполнении испытуемыми (т.е. не учитывая тех испытуемых, которые показали одинаковые результаты при выполнении каждого из этих заданий), можно объяснить на первый взгляд парадоксальную ситуацию, при которой у задания X_3 и оценка трудности, и оценка легкости оказались больше, чем соответствующие оценки у задания X_2. Задание X_3 не смогли выполнить испытуемые 7 и 11, чьи оценки способности в сумме оказались больше, чем оценки способности испытуемых 10 и 13, не выполнивших задание X_2; в результате оценка трудности задания X_3 оказалась больше. С другой стороны, задание X_3 смог выполнить испытуемый 13 с самой низкой оценкой неспособности (самый неподготовленный, справившийся всего с одним заданием теста), и испытуемый 10. С заданием X_2 справились испытуемые 7 и 11, сумма оценок неспособности которых меньше, чем у 13-го и 10-го, потому и оценка легкости задания X_2 оказалась меньше, чем у X_3.

Сумма оценок трудности и легкости задания, как правило, меньше 1, что принципиально отличает их от традиционных оценок – относительных частот верного выполнения заданий испытуемыми (см. формулу (2)). Эти оценки принципиально не зависят друг от друга, поскольку отражают разные характеристики заданий с точки зрения их выполнения испытуемыми.

6.3. Оценки латентных параметров испытуемых и заданий

Выше показано, что в результате решения рассмотренных задач (7) и (10) каждому испытуемому можно поставить две разные оценки: оценку p_i^* способности, выставляемую по выполненным им заданиям теста, и оценку q_i^* неспособности, выставляемую по невыполненным заданиям. Эти оценки характеризуют уровень и структуру знаний испытуемого.

При сравнении оценок двух испытуемых может оказаться, что у одного из них оценка способности больше, а оценка неспособности меньше, чем у другого. Эта ситуация понятна: первый испытуемый подготовлен лучше, чем второй. Но может случиться и так, что у первого испытуемого обе оценки, способности и неспособности, больше, чем у второго. В этом случае для оценки уровня подготовленности испытуемого необходимо использовать третью оценку, полученную из этих двух. Предлагается в качестве такой оценки использовать потенциал* $p^*{}_i : q^*{}_i$ испытуемого, вычисляемый как отношение оценок способности и неспособности, а в качестве оценки θ_i^* латентного параметра испытуемого – натуральный логарифм потенциала*:

$$\theta_i^* = \ln(p_i^* / q_i^*)\,. \tag{11}$$

Для характеристики трудности задания предлагается использовать потенциал* задания $t^*{}_j : s^*{}_j$ (отношение оценок трудности и легкости), а в качестве оценки β_j^* латентного параметра задания – натуральный логарифм потенциала* задания:

$$\beta_j^* = \ln(t_j^* / s_j^*)\,. \tag{12}$$

Оценки латентных параметров испытуемых и заданий, полученные по предлагаемой методике для рассматриваемой матрицы тестовых результатов, приведены в табл. 3 и 4. По сравнению с традиционным подходом, существенно изменилась оценка латентного параметра 2-го испытуемого. Судя по данной оценке, он подготовлен лучше всех испытуемых, участвовавших в тестировании, в том числе лучше, чем 1-й испытуемый, выполнивший наибольшее число заданий – 9. Дело в том, что 1-й испытуемый, получив большую оценку способности, чем 2-й, в то же время получил и большую по сравнению со 2-м оценку неспособности, не выполнив довольно легкое задание X_5 (см. табл. 4). Поэтому потенциал* 2-го испытуемого оказался немного больше, чем у 1-го, и соответственно оценка уровня подготовленности в логитах – выше.

Что касается испытуемого 5 с неправильной структурой знаний, то рекордно высокая оценка способности ему не помогла, поскольку высокой оказалась и оценка его неспособности. В итоге его оценка уровня подготовленности в логитах оказалась заметно меньше, чем у 4-го испытуемого, выполнившего такое же число заданий. Тем не менее, ранг 5-го испытуемого не изменился: его итоговая оценка оказалась больше, чем соответствующие оценки испытуемых, выполнивших меньшее число заданий (см. табл. 4).

Итоговые ранги заданий также не изменились, как и интервал изменения оценок их латентных параметров. Интервал изменения оценок латентных параметров испытуемых несколько увеличился за счет испытуемого 13, чья оценка заметно меньше, чем при традиционном подходе.

7. Решение задач для других матриц тестовых результатов

Рассмотренная выше матрица тестовых результатов представляет собой модельный пример, со специально подобранными свойствами. В табл. 5 показаны результаты применения традиционной и предлагаемой методики к другой, менее «правильной» матрице.

Если для матрицы из табл. 4 в результате применения предлагаемых оценок изменялся ранг (порядок) испытуемых, то для матрицы, представленной в табл. 5, изменяется ранг заданий. Самую низкую оценку латентного параметра имеет задание X_3, выполненное 8 испытуемыми (наиболее подготовленными; наименее подготовленные не справились с этим заданием). Несмотря на то, что с заданием X_1 справились 9 испытуемых, его оценка латентного параметра больше.

Для матрицы из табл. 5 в результате применения подхода, основанного на оценках способности и неспособности, увеличиваются по сравнению с традиционными оценками интервалы изменения латентных параметров как испытуемых, так и заданий.

В последних строках матрицы тестовых результатов из табл. 5 представлены результаты двух самых неподготовленных испытуемых, 12 и 13, выполнивших только 1 задание теста (соответственно X_4 и X_1). Их оценки латентных параметров близки. Возникает вопрос: а что изменится, если испытуемый 12 даст верный ответ не на задание X_4, а на самое трудное задание X_{10}? Ответ на этот вопрос дает табл. 6. Итоговая оценка в логитах 12-го испытуемого, выполнившего одно, зато самое трудное, задание теста (т.е. обладающего заведомо неправильной структурой знаний), оказалась заметно больше, чем у 13-го испытуемого, с более правильной структурой знаний, выполнившего одно из самых легких заданий теста. Тем не менее, ранг 12-го испытуемого не изменился: все испытуемые, выполнившие большее количество заданий теста, получили и большую по сравнению с ним оценку.

Сравнение табл. 5 и 6 приводит к еще одному наблюдению относительно разницы традиционного и предлагаемого подходов. При изменении профиля ответов одного испытуемого в традиционном подходе меняются только оценки этого испытуемого и измененных заданий. При подходе, основанном на оценках способности и неспособности, изменяются все без исключения оценки и испытуемых, и заданий.

Пример вычислений для еще одной матрицы приведен в табл. 7. Здесь, по сравнению с традиционным подходом, наблюдается значительная дифференциация по оценкам латентных параметров для заданий, с которыми справилось одинаковое число испытуемых (X_5 и X_6, X_7 и X_8), а также существенно большая оценка задания X_{10}. Изменяется также ранг испытуемого 2, а оценка самого неподготовленного испытуемого 13 существенно снижается.

7.1. Свойства предлагаемых оценок

Суммируя вышесказанное, можно перечислить следующие свойства и основные характеристики оценок, получаемых в результате решения задач об определении способности и неспособности (7) и (10):

1. Оценка p_i^* способности испытуемого выставляется с учетом трудности выполненных им заданий и учитывает уровень его подготовленности.
2. Оценка q_i^* неспособности испытуемого выставляется с учетом легкости невыполненных им заданий и учитывает как уровень подготовленности испытуемого, так и его структуру знаний.
3. Отсутствует однозначная зависимость оценок p_i^* способности и q_i^* неспособности испытуемого от набранного им исходного балла, а сумма этих оценок не равна 1.

4. Оценка t_j^* трудности задания выставляется с учетом способности справившихся с ним испытуемых и учитывает трудность этого задания для выполнения.
5. Оценка $s^*{}_j$ легкости задания выставляется с учетом неспособности выполнивших его испытуемых и учитывает легкость этого задания для выполнения.
6. Отсутствует однозначная зависимость оценок t_j^* трудности и $s^*{}_j$ легкости задания от исходного балла задания, а сумма этих оценок не равна 1.
7. Итоговая оценка θ_i^* (11) уровня подготовленности испытуемого в логитах вычисляется по его оценкам способности и неспособности и учитывает как уровень, так и структуру его знаний.
8. Итоговая оценка β_j^* (12) уровня трудности задания в логитах вычисляется по его оценкам трудности и легкости, т.е. оценкам как выполнимости, так и невыполнимости задания для испытуемых с учетом их уровня подготовленности.
9. Отсутствует однозначная зависимость оценок θ_i^* испытуемых и β_j^* заданий от исходных баллов, хотя порядок расположения (ранг) испытуемых и заданий в зависимости от их исходных баллов в основном сохраняется.

8. Выводы

В результате решения поставленных задач (7) и (10) каждый испытуемый получает две «педагогические» оценки, характеризующие как уровень его подготовленности, так и уровень его неподготовленности. Вычисленные на основе этих оценок оценки латентных параметров θ_i^* (11) и β_j^* (12), в отличие от традиционного подхода, основанного на долях верно и неверно выполненных заданий, учитывают не только количество полученных баллов, но и меру правильности структуры знаний испытуемых. Для матриц тестовых результатов со структурой, близкой к верхнетреугольной, полученные оценки не слишком отличаются от традиционных. Различия становятся более заметными по мере роста случаев нарушения правильности структуры знаний у испытуемых (См. далее табл. 5) .

Табл. 5. Матрица 2 тестовых результатов с оценками уровней подготовленности испытуемых и трудности заданий. Сравнение традиционного и предлагаемого подходов

	№№	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	R_i	W_i	p_i	p^*_i	q_i	q^*_i	p_i/q_i	p^*_i/q^*_i	$\ln(p_i/q_i)$	$\ln(p^*_i/q^*_i)$
	1	1	1	1	1	1	1	1	0	1	1	9	1	0,9	0,868	0,1	0,061	9	14,22	2,197	2,655
	2	0	1	1	1	1	1	1	1	1	0	8	2	0,8	0,739	0,2	0,203	4	3,640	1,386	1,292
	3	1	1	1	1	1	0	0	1	1	0	7	3	0,7	0,616	0,3	0,243	2,333	2,535	0,847	0,930
	4	1	1	1	0	0	1	0	1	0	1	6	4	0,6	0,592	0,4	0,376	1,5	1,574	0,406	0,454
	5	0	0	1	0	1	1	1	1	1	0	6	4	0,6	0,582	0,4	0,459	1,5	1,268	0,406	0,237
	6	1	0	1	1	1	0	1	0	0	0	5	5	0,5	0,407	0,5	0,369	1	1,103	0	0,098
	7	1	1	1	1	0	1	0	0	0	0	5	5	0,5	0,368	0,5	0,337	1	1,092	0	0,088
	8	1	1	1	1	0	0	0	0	0	0	4	6	0,4	0,275	0,6	0,443	0,667	0,621	-0,406	-0,477
	9	1	1	0	0	1	0	0	0	0	0	3	7	0,3	0,231	0,7	0,591	0,429	0,391	-0,847	-0,939
	10	1	1	0	0	0	0	1	0	0	0	3	7	0,3	0,25	0,7	0,567	0,429	0,441	-0,847	-0,819
	11	0	0	0	0	0	1	1	0	0	0	2	8	0,2	0,201	0,8	0,778	0,25	0,258	-1,386	-1,353
	12	0	0	0	1	0	0	0	0	0	0	1	9	0,1	0,092	0,9	0,879	0,111	0,105	-2,197	-2,257
	13	1	0	0	0	0	0	0	0	0	0	1	9	0,1	0,077	0,9	0,818	0,111	0,094	-2,197	-2,363
	R_j	9	8	8	7	6	6	6	4	4	2	Исходный балл		Доля верных ответов	**Способность**	Доля неверных ответов	**Неспособность**	Потенциал	**Потенциал***	Оценка ЛП	**Оценка ЛП***
	W_j	4	5	5	6	7	7	7	9	9	11										
Доля верных ответов	s_j	0,692	0,615	0,615	0,538	0,462	0,462	0,462	0,308	0,308	0,154										
Легкость*	s^*_j	0,621	0,461	0,407	0,414	0,315	0,362	0,398	0,209	0,158	0,071										
Доля неверн. ответов	t_j	0,308	0,385	0,385	0,462	0,538	0,538	0,538	0,692	0,692	0,846										
Трудность*	t^*_j	0,305	0,257	0,161	0,365	0,35	0,368	0,425	0,523	0,471	0,725										
Потенциал	t_j/s_j	0,444	0,625	0,625	0,857	1,167	1,167	1,167	2,25	2,25	5,5										
Потенциал*	t^*_j/s^*_j	0,491	0,558	0,396	0,882	1,111	1,017	1,068	2,502	2,981	10,21										

Оценка ЛП	$\ln(t_j/s_j)$	-0,811	-0,47	-0,47	-0,15	0,154	0,154	0,154	0,811	0,811	1,705										
Оценка* ЛП	$\ln(t^*_j/s^*_j)$	-0,711	-0,584	-0,927	-0,13	0,105	0,016	0,066	0,917	1,092	2,323										

Табл. 6. Матрица 3 тестовых результатов с оценками уровней подготовленности испытуемых и трудности заданий. Сравнение традиционного и предлагаемого подходов

	№№	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	R_i	W_i	p_i	p^*_i	q_i	q^*_i	p_i/q_i	p^*_i/q^*_i	$\ln(p_i/q_i)$	$\ln(p^*_i/q^*_i)$
	1	1	1	1	1	1	1	1	0	1	1	9	1	0,9	0,868	0,1	0,06	9	14,47	2,197	2,672
	2	0	1	1	1	1	1	1	1	1	0	8	2	0,8	0,748	0,2	0,241	4	3,104	1,386	1,133
	3	1	1	1	1	1	0	0	1	1	0	7	3	0,7	0,624	0,3	0,28	2,333	2,229	0,847	0,801
	4	1	1	1	0	0	1	0	1	0	1	6	4	0,6	0,587	0,4	0,34	1,5	1,726	0,406	0,546
	5	0	0	1	0	1	1	1	1	1	0	6	4	0,6	0,585	0,4	0,46	1,5	1,272	0,406	0,240
	6	1	0	1	1	1	0	1	0	0	0	5	5	0,5	0,415	0,5	0,407	1	1,020	0	0,019
	7	1	1	1	1	0	1	0	0	0	0	5	5	0,5	0,378	0,5	0,377	1	1,003	0	0,003
	8	1	1	1	1	0	0	0	0	0	0	4	6	0,4	0,284	0,6	0,481	0,667	0,590	-0,406	-0,527
	9	1	1	0	0	1	0	0	0	0	0	3	7	0,3	0,234	0,7	0,593	0,429	0,395	-0,847	-0,930
	10	1	1	0	0	0	0	1	0	0	0	3	7	0,3	0,252	0,7	0,571	0,429	0,441	-0,847	-0,818
	11	0	0	0	0	0	1	1	0	0	0	2	8	0,2	0,202	0,8	0,782	0,25	0,258	-1,386	-1,354
	12	0	0	0	0	0	0	0	0	0	1	1	9	0,1	0,174	0,9	0,939	0,111	0,185	-2,197	-1,686
	13	1	0	0	0	0	0	0	0	0	0	1	9	0,1	0,078	0,9	0,82	0,111	0,095	-2,197	-2,353
	R_j	9	8	8	6	6	6	6	4	4	3	Исходный балл		Доля верных ответов	**Способность**	Доля неверных ответов	**Неспособность**	Потенциал	**Потенциал***	Оценка ЛП	**Оценка ЛП***
	W_j	4	5	5	7	7	7	7	9	9	10										
Доля верных ответов	sj	0,692	0,615	0,615	0,462	0,462	0,462	0,462	0,308	0,308	0,231										
Легкость*	s^*_j	0,619	0,463	0,417	0,291	0,321	0,356	0,397	0,208	0,164	0,211										
Доля неверн. ответов	t_j	0,308	0,385	0,385	0,538	0,538	0,538	0,538	0,692	0,692	0,769										
Трудность*	t^*_j	0,315	0,268	0,173	0,389	0,36	0,38	0,435	0,531	0,48	0,7										
Потенциал	t_j/s_j	0,444	0,625	0,625	1,167	1,167	1,167	1,167	2,25	2,25	3,333										

Потенциал*	t^*_j/s^*_j	0,509	0,579	0,415	1,337	1,121	1,067	1,096	2,553	2,927	3,318										
Оценка ЛП	$\ln(t_j/s_j)$	-0,811	-0,47	-0,47	0,154	0,154	0,154	0,154	0,811	0,811	1,204										
Оценка* ЛП	$\ln(t^*_j/s^*_j)$	-0,676	-0,547	-0,88	0,29	0,115	0,065	0,091	0,937	1,074	1,199										

Табл. 7. Матрица 4 тестовых результатов с оценками уровней подготовленности испытуемых и трудности заданий. Сравнение традиционного и предлагаемого подходов

	№№	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	R_i	W_i	p_i	p^*_i	q_i	q^*_i	p_i/q_i	p^*_i/q^*_i	$\ln(p_i/q_i)$	$\ln(p^*_i/q^*_i)$
	1	1	1	1	1	1	0	1	1	1	1	9	1	0,9	0,866	0,1	0,105	9	8,248	2,197	2,110
	2	1	1	1	1	1	1	1	1	0	0	8	2	0,8	0,654	0,2	0,055	4	11,89	1,386	2,476
	3	1	1	1	1	0	1	0	1	0	0	6	4	0,6	0,442	0,4	0,229	1,5	1,930	0,406	0,658
	4	1	1	1	1	0	0	0	1	0	0	5	5	0,5	0,308	0,5	0,335	1	0,919	0	-0,084
	5	1	1	1	1	1	0	0	0	0	0	5	5	0,5	0,299	0,5	0,283	1	1,057	0	0,055
	6	1	1	1	0	0	1	0	0	0	0	4	6	0,4	0,262	0,6	0,351	0,667	0,746	-0,406	-0,292
	7	1	1	0	1	1	0	0	0	0	0	4	6	0,4	0,241	0,6	0,434	0,667	0,555	-0,406	-0,588
	8	0	0	0	0	1	0	1	0	1	0	3	7	0,3	0,374	0,7	0,776	0,429	0,482	-0,847	-0,730
	9	1	1	0	0	0	0	1	0	0	0	3	7	0,3	0,186	0,7	0,526	0,429	0,354	-0,847	-1,040
	10	1	1	0	0	0	1	0	0	0	0	3	7	0,3	0,204	0,7	0,502	0,429	0,406	-0,847	-0,900
	11	0	0	1	0	0	1	0	0	0	0	2	8	0,2	0,192	0,8	0,743	0,25	0,258	-1,386	-1,353
	12	1	0	1	0	0	0	0	0	0	0	2	8	0,2	0,09	0,8	0,613	0,25	0,147	-1,386	-1,919
	13	1	0	0	0	0	0	0	0	0	0	1	9	0,1	0,032	0,9	0,765	0,111	0,042	-2,197	-3,174
	R_j	11	9	8	6	5	5	4	4	2	1	Исходный балл		Доля верных	**Способность**	Доля неверных	**Неспособность**	Потенциал	**Потенциал***	Оценка ЛП	**Оценка ЛП***
	W_j	2	4	5	7	8	8	9	9	11	12										
Доля верных ответов	sj	0,846	0,6923	0,615	0,462	0,385	0,385	0,308	0,308	0,154	0,077										
Легкость*	s^*_j	0,734	0,493	0,475	0,252	0,289	0,329	0,256	0,127	0,154	0,018										
Доля неверн. ответов	t_j	0,154	0,308	0,385	0,538	0,615	0,615	0,692	0,692	0,846	0,923										
Трудность*	t^*_j	0,136	0,166	0,25	0,323	0,413	0,577	0,499	0,453	0,701	0,791										

Потенциал	t_j/s_j	0,182	0,444	0,625	1,167	1,6	1,6	2,25	2,25	5,5	12
Потенциал*	t^*_j/s^*_j	0,185	0,337	0,526	1,282	1,429	1,754	1,949	3,567	4,552	43,94
Оценка ЛП	$\ln(t_j/s_j)$	-1,705	-0,811	-0,47	0,154	0,47	0,47	0,811	0,811	1,705	2,485
Оценка* ЛП	$\ln(t^*_j/s^*_j)$	-1,686	-1,089	-0,642	0,248	0,357	0,562	0,667	1,272	1,516	3,783