

МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2024
MATHEMATICS AND EDUCATION IN MATHEMATICS, 2024
*Proceedings of the Fifty-Third Spring Conference
of the Union of Bulgarian Mathematicians
Borovets, April 1–5, 2024*

**POSSIBLE SCORING OF THE CONSTRUCTED-RESPONSE
TASKS FROM THE NATIONAL EXTERNAL ASSESSMENT
AFTER GRADE VII IN 2023, BASED ON ITEM RESPONSE
THEORY***

Pavlin Tsonev

Bulgarian Air Force Academy “Georgi Benkovski”,
Dolna Mitropoliya, Bulgaria
e-mail: ptsonev@yahoo.com

The article analyzes the constructed-response tasks from the National External Assessment after Grade 7 held in 2023. Scoring that is different from the generally accepted is justified.

Keywords: Item Response Theory, probabilistic modeling, constructed response tasks.

**ПРИМЕРНО ТОЧКУВАНЕ НА ЗАДАЧИТЕ СЪС
СВОБОДЕН ОТГОВОР ОТ НАЦИОНАЛНОТО ВЪНШНО
ОЦЕНЯВАНЕ СЛЕД VII КЛАС ПРЕЗ 2023 ГОДИНА,
БАЗИРАНО НА ВЕРОЯТНОСТНО МОДЕЛИРАНЕ**

Павлин Цонев

Военно-въздушно училище „Г. Бенковски“,
Долна Митрополия, България
e-mail: ptsonev@yahoo.com

Чрез методите на вероятностното моделиране са анализирани задачи със свободен отговор от НВО след VII клас проведено през 2023 г. Обоснован е алтернативен на общоприетия вариант за точкуване на такива задачи.

Ключови думи: Item Response Theory, вероятностно моделиране, тестови задачи със свободен отговор.

Предварителни бележки. Статията е продължение на изследванията на автора върху Националните външни оценявания (НВО) по математика след VII клас, проведени върху задачите с избираем отговор ([6] и [7]).

* <https://doi.org/10.55630/mem.2024.53.161-168>

Изследването е подкрепено от проект №80-10-61/25.4.2023 г. към Фонд научни изследвания на Софийски университет „Св. Климент Охридски“, 2023–2024 г.

2020 Mathematics Subject Classification: 97-11, 97D60.

Задачите със свободен отговор, които се използват в НВО по математика след VII клас, проведено през 2023 г., са два вида:

– с кратък свободен отговор, на които ученикът трябва да напише правилния според него отговор, който е едно или няколко числа;

– с разширен свободен отговор, на които ученикът трябва да напише подробно и логически обосновано решение, спазвайки методите, изучени в училище.

Употребата на комбинация от тези два вида задачи повтаря модела, който е наложен и при Национално външно оценяване от IV клас [3].

Точкуването на задачите със свободен отговор често е свързано с определени трудности. Две от тях са:

– Как да се определи максималният брой точки за дадена задача?

– Как да се разпределят тези точки върху решението на задачата?

В стремежа си оценяването да е максимално обективно често се дават точки, които освен цели числа, са и кратни на 0,5, дори и на 0,25.

Дали това е оправдано?

В настоящото изследване се дава една гледна точка на този въпрос, която се основава на методите на вероятностното моделиране (*Item Response Theory – IRT*) [4]. Тези методи за изследване на тестове все повече навлизат в България в последните години, например [1]. В теорията на вероятностното моделиране са разработени различни модели. Тук е използван еднопараметричният модел, наречен *Rash model*. За нуждите на IRT задачите с кратък свободен отговор трябва да се кодират с 0 точки за грешен отговор и 1 точка за верен отговор, а задачите с разширен свободен отговор да се кодират с цели числа в интервала $[0, n]$, като най-често $n = 2$ или $n = 3$. В статията е прието, ако една задача е кодирана с 0 или 1 точки, да се нарича *0-1 задача*, ако е кодирана с 0, 1 или 2 точки, да се нарича *0-1-2 задача* и т.н.

Основната идея в *Rash model* е, че възможността един ученик да реши дадена задача може да се измери чрез абстрактна величина, наречена *способност* (означена с θ). Вероятността този ученик да реши съответната задача се описва чрез т. нар. *характеристична функция* (графиката на която се нарича *характеристична крива*).

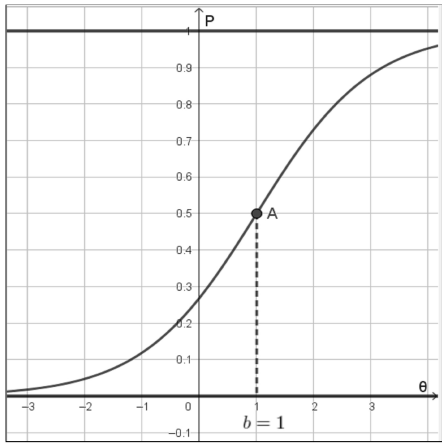
Ако задачата е 0-1, характеристичната функция има вида

$$P(\theta) = \frac{e^{\theta-b}}{1 + e^{\theta-b}},$$

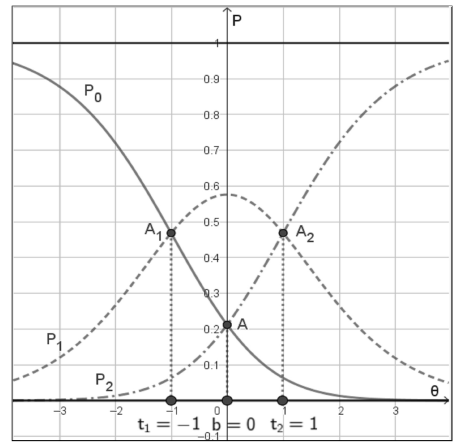
където b е параметър, който се нарича *трудност на задачата*. Графиката на такава крива при $b = 1$ е изобразена на Фигура 1а).

Координатите на точката от характеристичната крива с ордината $P = 0,5$ (в случая точка A) показват, че ученик със съответната способност (в случая $\theta = 1$) има вероятност 50% да получи верен отговор на тази задача. Тази стойност на θ (в случая $b = 1$) е трудността на задачата. Обикновено трудността на задачите е в интервала $(-2; 2)$. Ако b е близо до -2 , значи задачата е лесна, а ако е близо до 2 , значи задачата е трудна.

Ако задачата е 0-1-2, се чертаят три характеристични криви. На Фигура 1б) е даден пример, в който графиките P_0 , P_1 и P_2 съответстват на вероятността за получаване на 0, 1 и 2 точки. Трудността на такава задача се получава от абсцисата на пресечената точка A на P_0 и P_2 (в случая $b = 0$). Интерес представляват абсцисите



а)

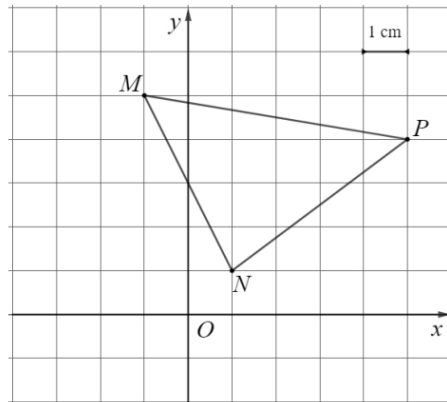


б)

Фиг. 1. Характеристични криви на Rash model

t_1 и t_2 на пресечните точки A_1 и A_2 съответно на P_0, P_1 и на P_1, P_2 (в случая $t_1 = -1$ и $t_2 = 1$). Изводът, който може да се направи при такава задача е, че учениците със способности по-малки от $t_1 = -1$ е най-вероятно да получат 0 точки, тези със способности между $t_1 = -1$ и $t_2 = 1$ е най-вероятно да получат 1 точка, а тези със способности над $t_2 = 1$ е най-вероятно да получат 2 точки. Проблемни задачи са такива, при които $t_1 > t_2$. По-нататък в статията са коментирани такива примери.

20. В координатна система Oxy с единична отсечка 1 cm са дадени точките M, N и P .



А) Запишете координатите на точките M, N и P .

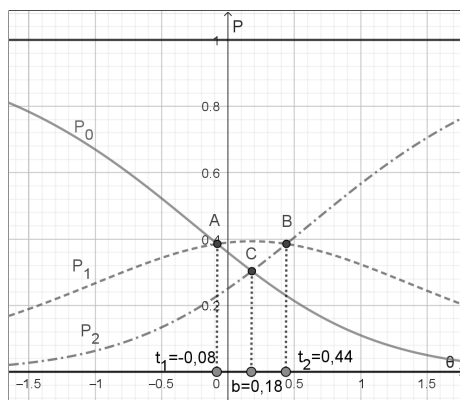
Б) Запишете координатите на точка Q , симетрична на точката N спрямо координатното начало.

Фиг. 2. Условие на задача 20 А) и Б)

Приложение на IRT при НВО след VII клас от 2023 г. Използвани са резултатите от НВО след VII клас, проведено през 2023 г. в 4 области на България – София-град, Пловдив, Плевен и Разград. Изследвани са получените данни от задачите със свободен отговор на 21 060 ученици. Данните са обработени със приложението jMetrik [5].

За да се приложи вероятностният метод, кодирането на задачите с кратък свободен отговор от 2023 г. се променя на 0-1. Такива са задачите 19 А), 19 Б) и 20 В), при които кодирането е било съответно 0-4, 0-3 и 0-4.

При някои задачи, като задача 20 Б) от Фигура 2, кодирането е 0-0,5-1 точки. За да се използва IRT методиката, кодирането се променя в цели числа на 0-1-2. Получените характеристични криви са изобразени на Фигура 3.



Фиг. 3. Характеристични криви на задача 20 Б)

Изводът, който може да се направи е, че учениците със способности под $-0,08$ е най-вероятно да получат 0 точки, тези със способности между $-0,08$ и $0,44$ е най-вероятно да получат 1 точка, а тези със способности над $0,44$ е най-вероятно да получат 2 точки. Тази задача има добри характеристики по отношение на трудност ($b = 0,18$) и по отношение на разслояването на учениците по трите точки (това се получава при $t_1 < t_2$).

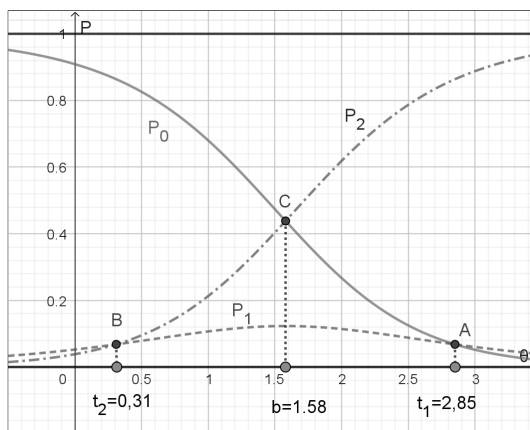
Пример за задача с $t_1 > t_2$ е задача 23 В), условието на която е дадено на Фигура 4.

23. В $\triangle ABC$ $AC = BC$, BL ($L \in AC$) е ъглополовящата на $\sphericalangle ABC$ и $\sphericalangle BLC = 60^\circ$. През средата на BL е построена права PQ , така че $PQ \perp BL$ ($P \in AB$, $Q \in BC$).

- А) Намерете ъглите на $\triangle ABC$.
- Б) Докажете, че $PBQL$ е ромб.
- В) Докажете, че $AL = BQ$.
- Г) Докажете, че $AP > PQ$.

Фиг. 4. Условие на задача 23

Оценяването на задачата е с 0,5 точки за доказано $LP = BQ$ и още 0,5 точки за доказано $AL = BQ$. Отново се приема точкуване 0-1-2, за нуждите на IRT. Така се получават графиките на характеристичните криви, изобразени на Фигура 5. При тази задача е най-вероятно да има два типа ученици – със способности под 1,58, които е най-вероятно да получат 0 точки, и такива със способности над 1,58, които е най-вероятно да получат 2 точки. С други думи, очаква се да има само два типа ученици – такива, които не са доказали нито едното равенство и такива, които са доказали и двете равенства. Ученици, които са доказали само едното равенство, се срещат с малка вероятност. Затова такава задача е подходящо да се кодира с 0 и 1 точки.



Фиг. 5. Характеристични криви на задача 23 В)

При други задачи с разширен свободен отговор кодирането е прекалено раздробено, за да може да се приложи IRT методиката. По тази причина е необходимо броят на точките в тяхното оценяване да е цяло число и да е сравнително малко. В [2] е анализиран начинът на точкуване на задачите, които оценяват частичния напредък на ученика при решаването им. Въвежда се понятието *важна стъпка*. Целта на важните стъпки е да се разграничат смислените моменти при решаването на дадена задача. Тези моменти се определят от автора на задачата. Пример за задача с разширен свободен отговор, която е разделена на голям брой важни стъпки, е задача 20 А) от Фигура 2. Оценяването на тази задача е през 0,25 единици. В таблица 1 са дадени точките и процента на получилите тези точки. Както се вижда, процентът на учениците получили някои точки е близко до 0%.

Таблица 1. Оценяване на задача 20 А) според МОН

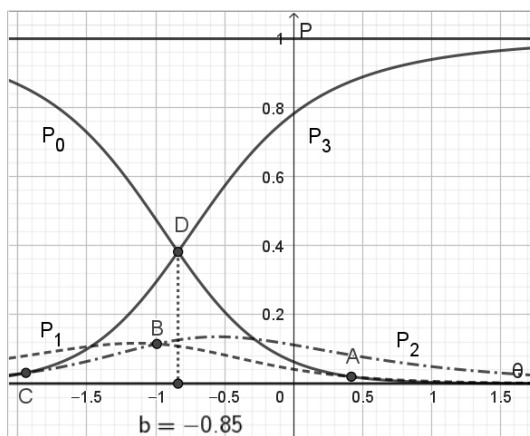
Точки	0	0,25	0,5	0,75	1	1,25	1,5	1,75	2	2,25	2,5	2,75	3
20 А)	34,7%	0%	1,6%	0%	3,9%	0%	2,6%	0%	4,1%	0%	6,4%	0%	46,6%

Това дава основание да се отделят 4 важни стъпки, които позволяват да се направи промяна в точкуването. Те са предложени в таблица 2. По този начин задачата става 0-1-2-3 и може да бъде изследвана с IRT. След като се извърши анализът на

данните на учениците, се получават характеристичните криви, изобразени на Фигура 6. От графиките се вижда, че учениците е най-вероятно да имат 0 или 3 точки, а междинните точки са малко вероятни. С други думи, ако един ученик може да намери координатите на една точка, той може да намери координатите на трите точки от условието на задачата. Тази задача също е пример за ситуация, при която $t_1 > t_2$. Изводът, който може да се направи е, че тя би трябвало да се кодира 0-1. Освен това, трудността на задачата е $b = -0,85$, което означава, че е от лесните задачи в теста.

Таблица 2. Ново точкуване на задача 20 А)

Точки от проверката	Обяснение	IRT
0-0,25		0
0,50-1,00	получени са 1 или 2 верни координати	1
1,25-2,00	получени са 3 или 4 верни координати	2
2,25-3,00	получени са 5 или 6 верни координати	3



Фиг. 6. Характеристични криви на задача 20 А)

По подобен начин са изследвани всички задачи със свободен отговор. След това задачите са подредени според коефициента им на трудност (b). Отчетени са графиките на характеристичните им криви, получени от IRT, и е коригиран броят на важните стъпки при решаването им. Оказва се, че се открояват най-много 2 важни стъпки. На тази база е създадено примерно ръководство за точкуване в таблица 3.

Както е известно, общият брой на точките на задачите със свободен отговор трябва да е 50. Затова, получените дотук точки на всяка задача се умножават с коефициент, показан в колона 5 на таблица 4. В случая този коефициент е съобразен с максималния брой точки, които са дадени от МОН на всяка задача (колона 3 на същата таблица). Освен това е взето предвид и нивото на трудност на задачите. По тази причина някои задачи имат разлики в точкуването в колоните 3 и 6. Например, задача 19Б има по-висок коефициент на трудност ($b = 1,68$) от задача 20А ($b = -0,85$). Затова точките на 19Б са увеличени от 3 на 4, а на 20А са намалени от 3 на 2.

Таблица 3. Примерно разпределение на точките за задачите със свободен отговор

№	Изпълнена дейност	Предложение
19А	получено 3	1
19Б	получено 10	1
20А	получени верни координати на поне 1 точка	1
20Б	<ul style="list-style-type: none"> • получена 1 вярна координата на Q • получена е втора вярна координата на Q 	<ul style="list-style-type: none"> • 1 • 1
20В	получено $S = 11$	1
21А	намерено разлагане и решено полученото уравнение	1
21Б	получени корени 3 и 9	1
21В	<ul style="list-style-type: none"> • сгрешено неравенство, но верни изводи за получените корени в получения от неравенството интервал • верни изводи, че 3, 6 и 9 са решения на неравенството 	<ul style="list-style-type: none"> • 1 • 1
22А	<ul style="list-style-type: none"> • въведени неизвестни x и $x - 2$ • получени брой работници 8 и 6 	<ul style="list-style-type: none"> • 1 • 1
22Б	намерени 5 и 4 дни	1
23А	<ul style="list-style-type: none"> • правдоподобен чертеж и намерен ъгъл 40° или 100° • намерени и трите ъгъла 40°, 40° и 100° 	<ul style="list-style-type: none"> • 1 • 1
23Б	<ul style="list-style-type: none"> • установено едно равенство, което доказва ромб • завършено е доказателството за ромб 	<ul style="list-style-type: none"> • 1 • 1
23В	доказано $AL = BQ$	1
23Г	доказано $AP > PQ$	1

Таблица 4. Оценяване на задачите със свободен отговор от 2023 г.

№ задача	Коефициент на трудност b	точки от МОН	Предложение за точкуване	Коефициент	Максимален брой точки
19А	1.74	4	0-1	4	4
19Б	1.68	3	0-1	4	4
20А	-0.85	3	0-1	2	2
20Б	0.18	1	0-1-2	1	2
20В	2.01	4	0-1	4	4
21А	1.25	6	0-1	4	4
21Б	-0.71	2	0-1	2	2
21В	-0.4	4	0-1-2	1,5	3
22А	0.72	5	0-1-2	2	4
22Б	1.36	6	0-1	4	4
23А	0.41	3	0-1-2	2	4
23Б	1.99	3	0-1-2	2	4
23В	1.58	2	0-1	4	4
23Г	2.79	4	0-1	5	5
общо точки		50			50

В заключение. Тази идея за оценяване на задачите със свободен отговор би могла да бъде приложена при бъдещи НВО. Така проверителите ще дават цяло число точки за всяка важна стъпка (не повече от 2-3 на задача). Тези стъпки се определят предварително от авторите на теста. След това екип от експерти ще направи анализ на получените резултати и според получените коефициенти на трудност ще бъдат дадени коефициенти за всяка задача, за да се получи необходимата сума от краен брой точки (в случая 50).

Съвременните технологии позволяват обработката на данните да става в рамките на няколко часа. Работата на експертите също не би трябвало да отнеме повече от 1 работен ден. Може да се предположи, че така ще се повиши обективността при оценяването на НВО.

ЛИТЕРАТУРА

- [1] R. M. ALASHKA. Two-factor Probabilistic Models in Education. *Mechanics Transport Communications. Academic journal*, **14**, no. 3/2, (2016), VII-7-VII-12 (in Bulgarian)
- [2] K. BANKOV. Scoring Tasks and Using Scales in Standardized Assessments. In: *Mathematics Education 75 Years of Mission and History*, University Publishing House “St. Kliment Ohridski”, (2023), 73–86 (in Bulgarian)
- [3] T. CHEHLAROVA. National External Assessment in Mathematics in 4th Class in 2022. *Mathematics & Informatics*, **65**, no. 4 (2022) (in Bulgarian)
- [4] R. HAMBELTON, H. SWAMINATHAN, H. ROGERS. *Fundamentals of Item Response Theory*. SAGE Publications, 1991.
- [5] J. P. Meyer. jMetrik. In: *Handbook of Item Response Theory*, pp. 557–566. Chapman and Hall/CRC, 2017.
- [6] P. Tsonev. Some Conclusions on the Results of the National External Assessment in Mathematics for 7th Grade. *Mathematics & Informatics*, **66**, no. 6 (2022), 578–601 (in Bulgarian)
- [7] P. Tsonev. Comparison of the Results Based on the Multiplechoice Items from the National External Assessment in Mathematics After Grade VII Held in 2020, 2021 and 2022. *Annual International Scientific Conference of VVVU “G. Benkovski”*, D. Mitropolia, 2023, 353–363 (in Bulgarian)