

ИЗПИТИ ЧРЕЗ ТЕСТОВЕ. ДАЛИ УЧЕНИКЪТ Е НАЛУЧКВАЛ ОТГОВОРА?

Емил Келеведжиев

Представено е изследване за анализиране на резултати от изпитване на ученици чрез тестове с избираем отговор. Използвани са данните от ежегодно провежданото се в България международно състезание по информатика и компютърна грамотност „Бобър“.

Ключови думи: Състезание „Бобър“, Тестове с избираем отговор.

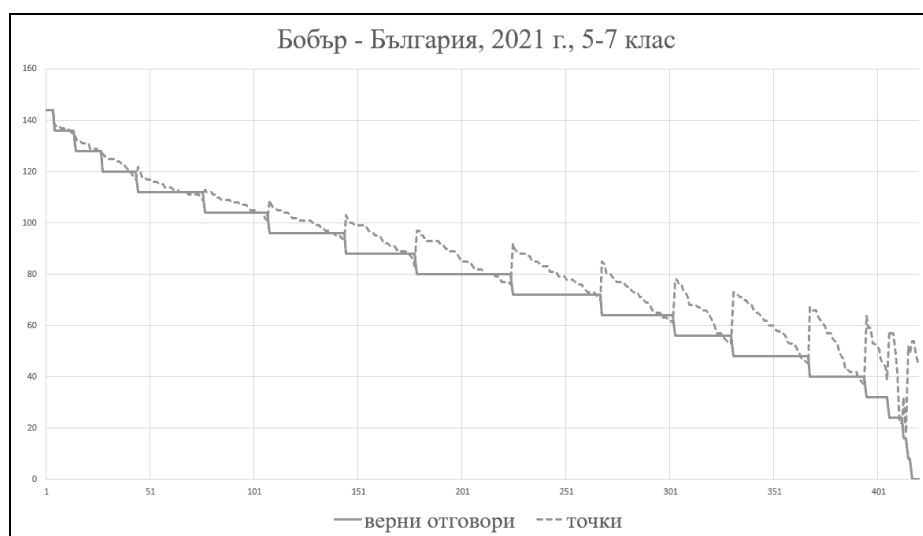
Увод. Дискусивно е доколко добре може да се извършва оценяване на знания чрез тестове с избираем отговор [1]. Даже има „ръководства“ [2] как да налучкваме отговорите. Става популярна тезата, че при изпитване чрез тестове, учениците в повечето случаи не решават поставената задача от теста, за да отбележат кой отговор са получили, а се опитват да го налучкат интуитивно дори случайно, или оглеждайки отговорите, да отхвърлят тези, които им изглеждат невъзможни. Това води до изкривяване на оценяването и оттам – до изкривяване на методите на обучение по предмети, в които уменията, а не запаметяването, имат основна роля – това са математиката и информатиката.

Разбира се, гореописаното е вярно при лошо направени тестове. За постигане на подобрение (понеже все пак не може да се откажем от тестовото оценяване поради ред причини, като предимството за обективност) предлагаме да се разнообразяват видовете тестове. Например да има тестове с повече от един верен отговор или отговорът да не се избира измежду дадени възможности, а да бъде отворен отговор, но все пак да позволява последваща автоматизирана обработка.

В състезанието „Бобър – България“ [3] (което е част от международното състезание „Бобър“ [4]) е приета система за оценяване чрез отнемане на точки за грешен отговор. Това дава възможност да се направи изследване доколко учениците са налучквали отговори и доколко реално са решавали задачите. Така може да се получи обща картина за цялата група ученици, а също и индивидуална преценка за всеки ученик. В следващото изложение ще използваме като синоними думите „въпрос“ и „задача“.

Постановка. В България състезанието „Бобър“ се провежда в две възрастови групи 5.–7. клас и 8.–9. клас. Предлагат се 18 въпроса с избираем отговор измежду 4 възможности, от които точно един е верен. Въпросите се задават на състезателя по реда на номерата им и не е възможно връщане към вече преминал въпрос. Въпросите са предварително разпределени по трудност в 3 групи, като във всяка от групите

те се оценяват с различен брой точки при верен отговор: За всеки верен отговор на въпрос с номер от 1 до 6 се дават по 4 точки, за верен отговор на въпрос с номер от 7 до 12 – по 5 точки и за верен отговор на въпрос с номер от 13 до 18 – по 6 точки. За грешен отговор на въпрос се отнемат по 3 точки. Когато състезателят остави въпрос без отговор, не се дават и не се отнемат точки за този въпрос. Крайният резултат на състезателя се получава, като към точките му се добави 54, за да няма отрицателни точки. Така максималният брой точки е 144. При състезатели с равен брой точки, по-напред в класиране се поставя състезателят, който е употребил по-малко сумарно време за решаване на задачите.

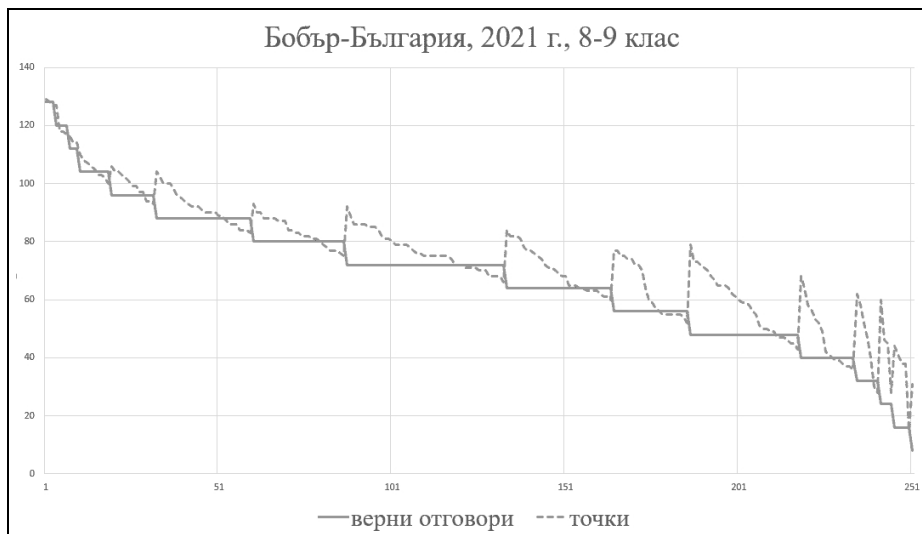


Фиг. 1. Графика за броя верни отговори и броя точки на участниците във възрастова група 5.-7. клас. Броят на верните отговори е нормиран чрез умножение с коефициент 8. По хоризонталната ос са дадени условните номера на учениците (вижте допълнителни пояснения в текста)

Задачите в състезанието се формулират кратко и включват елементи от всички области на информатиката: програмиране и алгоритми – структури от данни като пирамиди, стекове и опашки; моделиране на състояния, управление на процеси и данни; взаимодействието човек-компютър; графика и т. н. При умения на авторите за добро описание е възможно почти всички аспекти на компютърните науки и информационните технологии да бъдат тема на задача. Наборите от нови задачи и въпроси се създават и развиват в годишния международен семинар на състезанието.

Понеже състезанието се провежда изцяло компютъризирано, разполагаме с база от данни, върху която могат да се правят различни количествени и качествени изследвания. От многобройните публикации за такива изследвания може да посочим например [5] и [6].

Подход. Означаваме с t_i броя на верните отговори на i -тия ученик и с r_i – неговия резултат по точки, получени по правилата на състезанието. За сравнимост нормираме стойностите t_i , като вместо реалния брой, разглеждаме стойностите, ум-

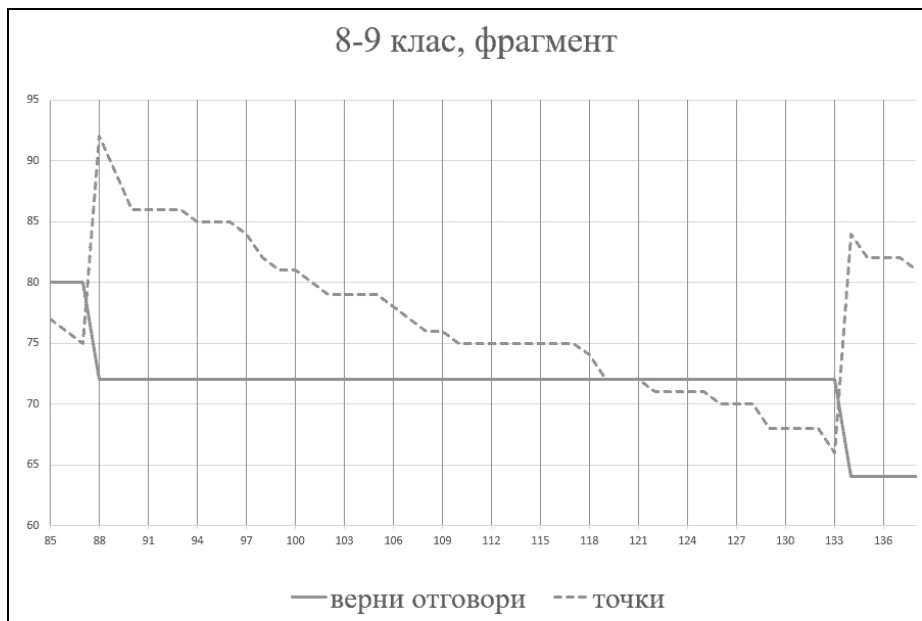


Фиг. 2. Графика за броя верни отговори и броя точки на участниците във възрастова група 8.–9. клас. Броят на верните отговори е нормиран чрез умножение с коефициент 8. По хоризонталната ос са дадени условните номера на учениците (вижте допълнителни пояснения в текста)

ножение по 8. Така при 18 въпроса, и двете величини t_i и r_i се изменят в диапазона от 0 до 144. За нагледност използваме условна номерация на учениците, подреждайки ги по намаляващ брой верни отговори, и във всяка група ученици с еднакъв брой верни отговори номерата на учениците са подредени по намаляващ брой точки, започвайки с ученика, получил най-много точки. Вижте фиг. 1 и фиг. 2, съответно за възрастови групи 5.–7. клас и 8.–9. клас от състезанието през 2021 г. Данните са взети от [3].

Във всяка група, състояща се от ученици с равен брой верни отговори, разглеждаме разликите $d_i = r_i - t_i$. За ученик, за когото d_i е близко до максималното в групата, е очевидно че този ученик е разглеждал задълбочено въпросите и не е отбелязвал произволен отговор, когато не е могъл да намери такъв. За ученик, за който тази разлика е отрицателна и близка до минимума в групата, следва че той е отбелязвал грешен отговор за по-голям брой въпроси от теста. Изводът е, че този ученик или е избирал случайно отговор, или даже разсъждавайки, не е могъл да реши задачата. И в двата случая това показва по-слаби знания, сравнено с учениците, за които d_i е близко до максималното в групата.

Този подход за допълнително оценяване на ученика (т.е. не само по точки) най-добре работи за групите ученици, наредени в средната част на класирането по брой верни отговори, което се вижда добре от графиките на фигурите. Ученик i с отрицателна стойност на d_i е по-склонен към случайно отгатване или има знания, покриващи само част от въпросите на теста. От гледна точка на съставителите на теста, наличието на много отрицателни стойности на d_i може да покаже, че въпросите не са подходящи за учениците.



Фиг. 3. Част от фиг. 2 в уголемен вид, съдържаща като пример точките на учениците с номера в интервала от 88 до 133. Всички те имат по 9 верни отговора ($72 = 9 \cdot 8$), но са получили различен брой точки. Учениците с номера в левия край на интервала $[88, 133]$ имат повече точки и са по-напред в класирането на състезанието. Това се дължи основно на факта, че те са се ограничили в посочването на отговори за тези от въпросите, за които не са сигурни, че могат да намерят верния отговор

Като пример на фиг. 3 е показан фрагмент от фиг. 2, където графично се виждат разликите d_i за учениците с номера i от интервала $[88, 133]$. В този интервал те имат еднакъв брой верни отговори. Там, където пунктираната линия е над плътната линия, съответните ученици имат положителни разлики d_i и са по-напред в класирането по точки.

Количествени индикатори. За целия тест дефинираме величините D и B чрез следните формули, където n е броят на учениците:

$$\text{Индекс на отрицателните разлики: } D = \frac{1}{n} \sum_{i=1}^n |\min\{0, d_i\}|$$

$$\text{Индекс на броя на отрицателните разлики: } B = \frac{\text{брой отрицателни стойности } d_i}{n}$$

Колкото по-големи са стойностите на D и B , с толкова по-голяма увереност може да предполагаме, че учениците са по-склонни към налучкване и/или имат знания, покриващи само част от въпросите. Пресметнати за данните от фиг. 1 и фиг. 2, индексите имат следната числова стойност:

$D \approx 0.23$ и $B \approx 0.12$ за възрастова група 5.–7. клас.
 $D \approx 0.66$ и $B \approx 0.28$ за възрастова група 8.–9. клас.

Заклучение. Сравнението между пресметнатите по-горе стойности на D и B за двете възрастови групи ни насочва към извода, че учениците от старшите класове са по-склонни към налучкване и/или имат знания, покриващи само част от въпросите за разглежданото състезание. Разбира се, следващи изследвания върху данните ще ни дадат по-точна преценка и уточняване на този извод, както и различни допълнителни интерпретации.

Предложеният подход би могъл да се използва за:

- обща качествена оценка за учениците, участващи в изпит чрез тестове;
- индивидуална качествена оценка за отделен ученик и за група ученици;
- анализиране на качествата на въпросите (задачите) в теста – дали авторите им са ги съставили подходящо за целите на изпита или състезанието.

ЛИТЕРАТУРА

- [1] Р. СТАНЧЕВА. Култура на оценяване с тестове в образованието. *Многообразие в единството*, изд. на Съюза на учените в България, кн. 2, 2010, 38–44. (Уеб страница: <http://www.tksi.org/SUB/papers/1-2/1-2-4.pdf> – посетена на 10.12.2021 г.).
- [2] Test-Taking Strategies for Multiple Choice Exams. Уеб страница: <https://lemongrad.com/test-taking-strategies-for-multiple-choice-exams/> (посетена на 10.12.2021 г.).
- [3] Бобър – България. Уеб сайт: <http://www.math.bas.bg/bbr> (посетен на 10.12.2021 г.).
- [4] Bebras International Challenge on Informatics and Computational Thinking. Уеб сайт: <http://bebras.org> (посетен на 10.12.2021 г.).
- [5] E. YAGUNOVA et al. Evaluation of Difficulty and Complexity of Tasks: Case Study of International On-line Competition „Beaver“. ISSEP'15, Slovenia (2015).
- [6] Е. КЕЛЕВЕДЖИЕВ, З. ДЖЕНКОВА. Количествени изследвания върху резултатите от състезанието „Бобър – България“. *Математика и математическо образование* **46** (2017), 246–251.

Емил Келеведжиев
Институт по математика и информатика
Българска академия на науките
ул. „Акад. Г. Бончев“ бл. 8
1113 София, България
e-mail: keleved@math.bas.bg

**MULTIPLE CHOICE QUESTIONS EXAM. HAS THE STUDENT
GUESSED THE ANSWER BY CHANCE?**

Emil Kelevedjiev

An analysis of students' multiple-choice questions exams is presented. The data of the "Bebras" international competitions in informatics and computer literacy held in Bulgaria are used.