Calculating the random guess scores of multiple-response and matching test items

Conference or Workshop Item

Calculating the random guess score of multiple-response and matching test items

Silvester Draaijer[a], Sally Jordan[b], Helen Ogden[c]

[a]VU University Amsterdam, Faculty of Behaviourals and Movement Sciences, Department of Research and Theory in Education, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands,
s.draaijer@vu.nl

[b]The Open University, School of Physical Sciences, Robert Hooke Building, Walton Hall, Milton Keynes, MK7 6AA, UK.
sally.jordan@open.ac.uk

[c]formerly Open University consultant and at the University of Southampton, UK.
h.e.ogden@southampton.ac.uk

Abstract

Abstract. For achievement tests, the guess score is often used as a baseline for the lowest possible grade for score to grade transformations and setting the cut scores. For test item types such as multiple-response, matching and drag-and-drop, determining the guess score requires more elaborate calculations than the more straightforward calculation of the guess score for True-False and multiple-choice test item formats. For various variants of multiple-response and matching types with respect to dichotomous and polytomous scoring, methods for determining the guess score are presented and illustrated with practical applications. The implications for theory and practice are discussed.

Keywords: item-writing, question development, test development, cut score, standard setting, selected response test items, selected response questions, multiple-choice questions, MCQs

adfa, p. 1, 2011.

1 Background

An essential step in test construction are the rules for setting cut score and score-to-grades. An important consideration in this step is determining the lower bound of the score range that students can achieve on the basis of random guessing. In this paper, methods and tables for calculating or looking up guess values for multiple-response and matching test items are presented and discussed.

First of all, there is not one ‘optimal’ method for establishing cut score [1–4]. A suitable method depends on the goal of the test and available resources. The main methods for standard setting can be classified as criterion referenced methods (setting a cut score on the basis of the content of the test and considerations of minimum levels of achievement needed related to that content), norm referenced methods (setting a cut score in relation to the score distribution of the population that took the test), and combining these methods somehow (setting a cut score based on a combination of both approaches). In many situations in higher education in the Netherlands and the UK, the random guess score for a test with selected response test items is taken into account [5–9] for both types of standard setting. The random guess score provides a criterion for the lowest score that can be awarded the lowest possible grade for a student. The assumption is that this is the score that is obtained by simply filling in answers randomly[1] but according to instructions for filling in (e.g. the instructions regarding the number of options to select for a test item).

With the advent and increased use of e-assessment [12–14], teachers in higher education can more easily than ever use test items types other than True-False or multiple-choice. In particular, multiple-response, matching and drag-and-drop test items can be deployed easily. The question therefore becomes more pressing how the guess score must be calculated for such items [15]. McKenzie & O’Hare [16] discussed the problems associated with establishing such a base guess factor for complex test item formats such as multiple-response and drag-and-drop questions. They argued that in the random response mode for such questions, nodes appear for groups of test-takers that achieve a certain score based on specific settings of question answering and that the guess factor is often more prominent than one would expect. They reported these findings on the basis of simulations they performed using a Marking Simulator. Unfortunately, since the publication of MacKenzie & O’Hare, no progress has been reported concerning the development of the Marking Simulator application. Further Jordan [17] presented a general approach to establishing guess values for multiple-response items, multiple attempt multiple-response items and drag-and-drop items. Her approach was very principled from a mathematical viewpoint and a stand-alone program was developed for use by experts. This leaves teachers in higher education and less mathematically proficient test item authors on their own in dealing with this problem.

---

[1] Some other methods try to incorporate student’s knowledge level in estimating guessing level using formula scoring [10] but this is abandoned because of validity problems [11].

In this article, we will put forward some methods and tables that allow testing experts and teachers in higher education to calculate or find the random guess score for multiple-response and matching type questions based on various set-ups of these items.

2 Basic principles for calculating the random guess score

In principle, the random guess score of a test item S_{guess}, equals the sum of the probability for each possible outcome for a question $p(O_i)$, multiplied by the score for that outcome O_i: S_{O_i}. This can be written:

$$S_{guess} = \sum p(O_i) * S_{O_i} \tag{1}$$

For True-False test-items, a dichotomous item, there are two combinations of choices possible. One combination leads to score 0 and one combination to the maximum score. The probability p of scoring 0 points is the number of occurrences of 0 points, divided by the total number of combinations. This is expressed as a probability $p = \frac{1}{2}$. Given a maximum score of 1 point, this random guess score is $S_{guess} = p(O_0) * 0 + p(O_1) * 1 = \frac{1}{2} * 0 + \frac{1}{2} * 1 = 0.5$ points.

For a four choice multiple-choice item, four combinations are possible of which three options lead to a score of 0 points and one leads to a score of 1 point. The random guess score now equals $S_{guess} = p(O_0) * 0 + p(O_1) * 1 = \frac{3}{4} * 0 + \frac{1}{4} * 1 = 0.25$ points.

3 Multiple-Response test items

A multiple-response test item is similar to a multiple-choice test item, but there is more than one correct answer. Multiple True-False test items are similar to multiple-response test item with regard to random guess score. For the random guess score of multiple response test item two characteristics are of importance.

1. Is the scoring of the test items dichotomous (correct or incorrect) or polytomous (multiple points can be acquired by specific selection of options)?
2. Is the examinee informed what the number of correct alternatives is?

3.1 Dichotomous scoring

For example, let us take a 5 alternative multiple-response test items of which three alternatives are correct. The student is instructed to select the three correct alternatives (out of five possible alternatives). Suppose we use a dichotomous scoring model in which the student receives 1 point if the answer is completely correct and 0 points for all other situations. For this test item we can calculate the number of combinations of possible choices as being $\binom{n}{m} = \frac{n!}{m!(m-n)!}$ which yields for this example $\binom{5}{3} = \frac{5!}{3!(5-3)!} = 10$.

Only 1 of those combinations leads to a score of 1 point, the rest leads to a score of 0 points. We can represent a specific combinations of choices as CO_i. Now, the random guess score can be calculated as follows: $S_{guess} = p(CO_{0pt}) * 0 + p(CO_{1pt}) * 1 = \frac{9}{10} * 0 + \frac{1}{10} * 1 = 0.1$ points.

3.2 Polytomous scoring

A different situation occurs if we use a polytomous scoring model for the test items in which the student receives 1 point for each correctly chosen alternative and the student is also *instructed* to select the three correct alternatives. For this test item, 10 combinations of selections are possible. By tabulating all possible options and assigning scores to each option, the random guess score can be calculated.
S_{guess} now follows from: $S_{guess} = p(CO_{0pt}) * 0 + p(CO_{1pt}) * 1 + p(CO_{2pt}) * 2 + p(CO_{3pt}) * 3 = 0 * 0 + \frac{3}{10} * 1 + \frac{6}{10} * 2 + \frac{1}{10} * 3 = 1.80$ points.

A more elegant approach to this calculation is given by Jordan [17] and a simpler form of that follows now. For the situation above, where the student is told in advance how many correct alternatives there are, we can find S_{guess} using a simple formula, which may be derived as follows: think of all the responses as balls in a bag. There are n "correct" balls in the bag, with labels $C_1, \dots, C_n$ on them, and m balls in total. The student is told to select n balls from the bag. Any one of the m balls is equally likely to be in the students' selection (therefore it can be regarded a random variable), with probability $\frac{n}{m}$, since there are m balls in total, and we pick n of them. So, the probability the first correct ball, C_1, is selected is $\frac{n}{m}$, the probability C_2, is selected is $\frac{n}{m}$, and so on.

A theorem in probability theory is that the expected value of the sum equals the sum of the expected values of the accompanying random variables, whether they are dependent or not. So we can write $S_{guess} = \sum p(O_i) * S_{O_i} = \sum \frac{n}{m} * S_{O_i}$

If we have the simple scoring rule where each correct response scores 1 point, then:

$$S_{guess} = \sum \frac{n}{m} * S_{O_i} = \frac{n^2}{m} \quad (2)$$

Or, we can write this as a percentage of the total possible achievable score of n points as $100 * \frac{n}{m}$%. If we apply this formula to the example above, where we are told that there are $n = 3$ correct answers of the are $m = 5$ total answers, $S_{guess} = \frac{n^2}{m} = \frac{9}{5} = 1.8$ points. So we arrive at the same result as we did by counting the possible combinations. If the students are told the number of correct responses, then we can extend the argument above to give:

$$S_{guess} = \sum \frac{n}{m} * S_{O_i} \tag{3}$$

where the sum is over all m possible choices, and where the "score" given for selecting an incorrect response may be actually be negative, to give a penalty for incorrect responses.

3.3 Giving the number of correct responses

It seems that in the case where we are given the number of correct responses, the random guess score should be fairly easy to find. However, if we are not given the number of correct responses, we could compute the random guess score by assuming that each possible selection of options is equally likely to be chosen. We will assume a random selection of options in the section below. Given this method, Table 1 is constructed. The table contains the random guess score for commonly encountered multiple-response test items. For multiple-response test items with different scoring rules, different tables should be constructed.

Table 1. Random guess scores for Multiple Response test items given the number of alternatives and number of correct alternatives.

Number of alternatives	Number of correct alternatives	Dichotomous (0 or 1)		Polytomous (each correct alternative 1 point)		
		Unknown number correct	*Known* number correct		*Unknown* number correct	*Known* number correct
n	*m*	Random guess score	Random guess score	Max score of item	Random guess score	Random guess score
3	1	0.13	0.33	1	0.5	0.33
3	2	0.13	0.33	2	1.0	1.33
3	3	0.13	1	3	1.5	3.00
4	1	0.06	0.25	1	0.5	0.25
4	2	0.06	0.17	2	1.0	1.00
4	3	0.06	0.25	3	1.5	2.25
4	4	0.06	1	4	2.0	4.00
5	1	0.03	0.2	1	0.5	0.20
5	2	0.03	0.1	2	1.0	0.80
5	3	0.03	0.1	3	1.5	1.80
5	4	0.03	0.2	4	2.0	3.20
5	5	0.03	1	5	2.5	5.00
6	1	0.02	0.17	1	0.5	0.17
6	2	0.02	0.07	2	1.0	0.67
6	3	0.02	0.05	3	1.5	1.50
6	4	0.02	0.07	4	2.0	2.67
6	5	0.02	0.17	5	2.5	4.17
6	6	0.02	1	6	3.0	6.00

As an example, consider the question shown in Figure 1. This test item contains 5 alternatives of which 2 are correct alternatives. Students are told the number of correct alternatives. If the scoring is dichotomous, Table 1 shows that the random guess score equals 0.1 points; if the scoring is polytomous, the random guess score equals 0.8

Figure 1

Fig. 2. Example random guess scores for a multiple response item.

A 45 year old asthmatic woman who has lived all her life in Glasgow presents with a goitre of four years' duration and clinical features suggestive of hypothyroidism. The *<u>two</u>* *most likely diagnoses include*

A. Iodine deficiency
B. Dyshormonogenesis
C. Drug-induced goitre
D. Thyroid cancer
E. Auto immune thyroiditis

Correct answer: true C and E: false A, B and D [18]

3.4 An extension for scoring rules for multiple-response test items

In specific circumstances, more sophisticated scoring might be required for a multiple-response test item. For the example given in Figure 1, the scoring rule could for example be defined as follows:

- 0 points: If the student gets 0 alternatives correct and 3 incorrect OR If the student gets 1 alternative correct and 2 incorrect
- 5 points: If the student gets 2 alternatives correct and 1 incorrect
- 10 points: If the student gets all 3 alternatives correct

Neither Equation (3) nor the more straightforward calculation table will now suffice. We can return to handwork and develop a new table with combinations, as shown in Table 2. This multiple-response test item can have 20 combinations. We must assign scores to each combination of choices. Then we can calculate the probability of occurrence of each score. The occurrence of the full score is 1/20th, the occurrence of a score of 5 points is 9/20th and the score of 0 points is 10/20th. From this, it follows that $S_{guess} = p(CO_{0pt}) * 0 + p(CO_{5pt}) * 5 + p(CO_{10pt}) * 10 = 2.75$ points. From the table, this can also be calculated by averaging the sum of scores.

Table 2. Combination Table for a Multiple-response Test item with 6 Options and 3 Correct alternatives, scores and average score.

	Correct or Wrong						
Combination	C	C	C	W	W	W	Score
1	x	x	x				10
2	x	x		x			5
3	x	x			x		5
4	x	x				x	5
5	x		x	x			5
6	x			x	x		0
7	x				x	x	0
8	x		x		x		5
9	x		x			x	5
10	x			x		x	0
11		x	x	x			5
12		x		x	x		0
13		x			x	x	0
14		x	x		x		5
15		x	x			x	5
16		x			x	x	0
17			x	x	x		0
18			x		x	x	0
19			x	x		x	0
20				x	x	x	0
Average score							2.75

A note of warning must be given here. The assumption that answers are selected completely at random is not likely to be realistic in practice; the answering behavior of students might play a role. A student is likely to make some guess as to how many of the answers they think will be correct, probably based on their past experience of answering test items of a similar type. Even in multiple-choice test items, guessing behavior is influenced by student characteristics, with for example students being more likely to select the inner options of a multiple-choice test item than the first and last alternative [19], as well as the quality of the test item and its foils.

For multiple-response test items, it would be interesting to see how many choices real students do assume to be correct (before they even look at the content of those choices) when answering this sort of test item. Once the distribution of the number of choices a student would assume to be correct, we could make a better substantiated calculation to find the random guess score.

4 Ordering and Matching test items

It is easier to compute the random guess scores of Ordering and Matching test item types than it is for multiple-response test items. An example of a matching item is shown in Figure 2.

Fig. 3. Example Matching Test item with 3 options and 4 markers [20].

Match the type of quiz question on the right with the correct description of it on the left. You can use the type of quiz only once.

______ *Students must make associations between items on two lists*	*A. Essay*
______ *Students judge the correctness of declarative propositions*	*B. Matching*
______ *Students choose one correct response from a list of options*	*C. Multiple-choice*
	D. True-False

It is interesting to note that drag-and-drop test items for which a student needs to place specific objects (for example text markers) in the correct boxes is also a matching test item. See the example of Figure 3. For the random guess score of matching test items, two characteristics are of importance.

1. Is the test item scoring dichotomous (correct or incorrect) or polytomous (multiple points can be acquired for each correct choice)
2. Can the answering options be used more than once or only once? For Ordering test items, the options (being ordering numbers) can only be used once. For matching and drag-and-drop this must follow from the specific set-up of the test item. In what follows, we will assume that the answering options can only be used once.

4.1 Dichotomous scoring

Let us assume a student has to answer a test item in which he has to position 5 answering options in 4 open spaces (see Figure 3). One answering option is a foil. As the test item is dichotomous, the student receives 1 point if all four answering options are set correct and the foil is left unused.

This problem can be approached by the analogy of marbles in a bag. In this case, there are 4 bags and 5 colored marbles that have to be put in the correct bag. This is drawing problem without replacement. The number of permutations for this problem is 5! which equals 120. The chance to score 1 point for this test item (all options correct) is 1 divided by the number of possible permutations which yields 0.0083. This equals the approach in which the chance the get the first item correct is 1/5, the second 1/4 and so forth, which yields $\frac{1}{5} * \frac{1}{4} * \frac{1}{3} * \frac{1}{2} * \frac{1}{1} = 0.0083$ points.

4.2 Polytomous scoring

Let us assume again that a student has to answer a test item in which he has to position 5 answering options in 4 open spaces. One answering option is a foil. Also suppose the student receives 1 point for each correct positioned answering option (which constitutes a correct match). In contrast to the dichotomous scoring, the calculation of the random guess score is more complicated. A direct approach to the problem would be to use the analogy of constructing a sequence of 4 numbers using the numbers 1 to 5 (1, 2, 3, 4 and 5) and establish how many permutations there are with the sequence 0123 and 4 on their correct positions. We could develop a table, but it would comprise 120 rows with unique sequences. Therefore, we could follow the reasoning described as follows.

- The number of permutations with all numbers on their correct position is 1.
- The number of permutations with only three numbers on their correct position is 4. These permutations are 1235, 1254, 1534, 5234
- The number of permutations with two numbers on their correct position is 18 because there are 6 possibilities to select 2 numbers from 1 to 4: 12, 13, 14, 23, 24, 34
  - When one starts for example to put the numbers 1 and 2 on the correct position, then 3 possibilities remain to put two incorrect numbers on their position (43, 53 and 45)
  - This line of reasoning also applies to the other 5 combinations of 2 correct positioned numbers.
- The number of permutations for which 1 number is positioned on its correct position is 44 because there are 4 possibilities to draw 1 number from the numbers 1 to 4
  - If we put for example number 1 on its correct position we only have 2 possibilities to position the number 234 incorrectly (432, 342). If we incorporate the number 5 in these sequences, 3 extra sequences will comply with the number 5 on position 4, 3 and 2 resulting in 9 sequences (325, 425,345, 352, 452, 453, 543, 542, 523). Therefore 11 sequences.
- The number of permutations for which not a single number is on its correct position is 53. For the number 1234 there are 9 possibilities and for the numbers 1235, 1254, 1534 and 5123 there are each 11 possibilities.
  - For numbers 1234: Choose in first instance number 2. Numbers 134 must be positioned incorrectly. There are 3 possibilities for that. The same counts when choosing number 3 or 4 on the first position.
  - For numbers 1235, 1254, 1534 and 5123 the same procedure applies which results per number combination in 11 possibilities. This gives a total of 3*3+4*11=53 permutations.

The expected random guess score now follows from: $S_{guess} = 4 * \frac{1}{120} + 3 * \frac{4}{120} + 2 * \frac{18}{120} + 1 * \frac{44}{120} + 0 * \frac{53}{120} = \frac{96}{120} = 0.8$ points

As can be seen, this approach is quite elaborate and can easily lead to calculation mistakes. A more elegant approach is the following. Suppose C_1 is a random variable that can have value 1 if number 1 is positioned on its correct position (first place) and a value of 0 if not correctly positioned. Define the random variable C_1 to C_n in the same way. The total score for a test item is defined as the sum of these random variables. A theorem in probability theory is – as we used with multiple-response test item guess score calculation - that the expected value of the sum equals the sum of the expected values of the accompanying random variables, whether they are dependent or not. Now suppose we have a matching test item with m markers that have to be matched with n options in which $n \leq m$. It then follows that $S_{guess} = \sum p(O_i) * S_{O_i}$ can be written as

$$S_{guess} = \sum \frac{1}{m} * S_{O_i} \tag{3}$$

If we apply this to the example above, the probability of having a value of 1 is $\frac{1}{5}$ and the probability of having value 0 is $\frac{4}{5}$ for each response. For each random variable, the expected value is $\frac{1}{m} * 1 = \frac{1}{5} * 1 = 0.2$ points and therefore the total expected value is 0.8. Given these calculations, Table 3 is constructed which displays the random guess score for common encountered matching test items.

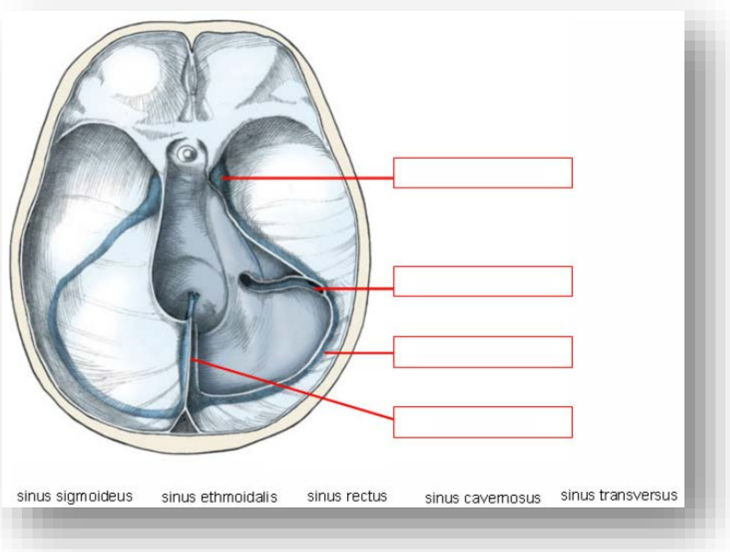
Table 3. Random Guess Scores for Matching Test items

Number of alternatives	Total number of match alternatives	Dichotomous (0 or 1 points)	Polytomous (each correct alternative 1 point)	
n	*m*	Random guess score	Max score of item	Random guess score
2	2	0.50	2	1.00
2	3	0.17	2	0.67
2	4	0.04	2	0.50
3	3	0.17	3	1.00
3	4	0.04	3	0.75
3	5	0.01	3	0.60
4	4	0.04	4	1.00
4	5	0.01	4	0.80
4	6	0.00	4	0.67
5	5	0.01	5	1.00
5	6	0.00	5	0.83
5	7	0.00	5	0.71
6	6	0.00	6	1.00
6	7	0.00	6	0.86
6	8	0.00	6	0.75
7	7	0.00	7	1.00
7	8	0.00	7	0.88
7	9	0.00	7	0.78

The test item shown in Figure 4 contains 4 alternatives and 4 matching items and 1 extra foil item. If the test item has dichotomous scoring, Table 3 shows that the random guess score equals 0.15 points If the test item has polytomous scoring, the random guess score equals 0.80 points

Fig. 4. Example random guess scores for a matching test item.

Drag the given terms at the bottom of the image to the corresponding boxes. <u>Note</u> that there is one foil term that must not be used



Discussion and Conclusion

For the purpose of establishing cut-scores and score-to-grade calculations for achievement tests, we have shown how to calculate guess values for a range of multiple-response and matching test items. Such calculations can be prone to calculation mistakes. We provided simple tables to look up random guess values for often used variants of these test items. These tables may prove their worth in the praxis of higher education for teachers and examiners using such item types in their assessments. These tables may prevent teachers from making calculation mistakes if they were to establish random guess values for themselves.

However, other more sophisticated approaches may be preferable. For examples computer tools that can work out the random guess score might be helpful. For example, platforms such as R in combination with online presentation and manipulation using shiny (https://www.rstudio.com/products/shiny/shiny-user-showcase/) could be used to make a friendly user interface and provide easy access to additional forms of scoring such as negative scoring, scoring with ceilings or using the so called ‘quotient rule’ by Vos et al. [21, 22]. Even more helpful could be if e-assessment tools would automatically provide the user with the random guess value. It is a matter of discussion for scholars, practitioners and vendors of e-assessment software at conferences such as the TEA to establish whether this would be an interesting line of research and development.

With respect to the findings of the random guess values themselves, we note that some items have maybe unexpectedly very high guess values. In particular polytomous scoring multiple-response items can have high guess values when the number of correct alternatives is given. It can be argued that these items should not be used in summative tests because they introduce a lot of error in the measurement. In fact, for optimal discrimination purposes, it is important to try to design test items that have about 50% chance of being answered correctly after deduction of the guess value [23]. The higher the guess value of a multiple-response test item, the smaller the interval remains in which discrimination of the test items will be able to be reached. Very low random guess values on the other hand, as with dichotomous scoring multiple-response and matching test items, can cause students with a bit less than perfect knowledge gain no points. In that situtation, items do not discriminate well either either. It requires careful consideration concerning the level of difficulty of the subject matter and estimations of the level of knowledge and skill of the student population to establish how multiple-response and matching test items should be designed and set up.

With respect to future research, studies investigating student preferences for specific positions of alternatives in multiple-choice test items [19], could be conducted. This study has noted that the expectations that students have regarding the correct number of alternatives for multiple-response test items (if the number of correct alternatives is not given) and the position these alternatives have, can be significant. Further work in this area could yield important additional information and design considerations for multiple-response test items and their application in achievement testing and other testing programs..

Acknowledgment

We would like to thank to Dick Neeleman of the Vrije Universiteit Amsterdam for his contribution to this article.

References

1. Berk, R.A.: A consumer’s guide to setting performance standards on criterion-referenced tests. Rev. Educ. Res. 56, 137–172 (1986).
2. Brennan, R.L.: Educational measurement. Rowman & Littlefield Publishers (2006).
3. Cizek, G.J.: Setting performance standards: concepts, methods, and perspectives. Routledge (2001).
4. Cohen-Schotanus, J., Van der Vleuten, C.P.M.: A standard setting method with the best performing students as point of reference: Practical and affordable. Med. Teach. 32, 154–160 (2010).

5. Burton, R.F.: Quantifying the effects of chance in multiple choice and true/false tests: Question selection and guessing of answers. Assess. Eval. High. Educ. 26, 41–50 (2001).
6. Burton, R.F.: On guessing corrections. Med. Educ. 38, 113–113 (2004).
7. Ebel, R.L., Frisbie, D.A.: Essentials of educational measurement. Prentice Hall, New Jersey (1991).
8. Gronlund, N.E.: Assessment of student achievement. Allyn & Bacon, Boston, MA (1998).
9. Van Berkel, H.J.M., Bax, A.: Toetsen in het Hoger Onderwijs [Testing in Higher Education]. Bohn Stafleu Van Loghum, Houten/Diegem (2006).
10. Frary, R.B.: Formula scoring of multiple-choice tests (correction for guessing). In: ITEMS: The Instructional Topics in Educational Measurement Series (1988).
11. Roberts, D.: Let’s talk about the “correction for guessing” formula. Penn State University (2001).
12. Draaijer, S., Hartog, R., Hofstee, J.: Guidelines for the design of digital closed questions for assessment and learning in higher education. E-J. Instr. Sci. Technol. 10, (2007).
13. Hartog, R., Draaijer, S., Rietveld, L.C.: Practical aspects of task allocation in design and development of digital closed questions in higher education. Pract. Assess. Res. Eval. 13, 2–15 (2008).
14. Jordan, S.: E-assessment: Past, present and future. New Dir. 9, 87–106 (2013).
15. Draaijer, S., Hartog, R.: Design patterns for digital item types in higher education. E-J. Instr. Sci. Technol. 10, (2007).
16. Mackenzie, D., O’Hare, D.: Empirical prediction of the measurement scale and base level “guess factor” for advanced computer-based assessments. In: Danson, M. (ed.) Sixth International Computer Assisted Assessment (CAA) Conference Proceedings. , Loughborough University (2002).
17. Jordan, H.: Random guess score, http://www.open.ac.uk/blogs/SallyJordan/wp-content/uploads/2011/05/RGS.pdf, (2009).
18. Brown, G.A., Bull, J., Pendlebury, M.: Assessing student learning in higher education. Routledge (1997).
19. Attali, Y., Bar-Hillel, M.: Guess where: The position of correct answers in multiple-choice test items as a psychometric variable. J. Educ. Meas. 40, 109–128 (2003).
20. Kupsch, B., Horn, E.: Writing matching questions. UW–Madison School of Nursing.
21. Vos, H., De Graaf, A.: De quotiëntregel, (2008).
22. Vos, H., Kloppenburg, M., Tomson, O.: Optimale uniforme scoringsregels voor innovatieve vraagvormen [Optimal uniform scoring rules for innovative test item formats], (2010).
23. Eggen, T.J., Lampe, T.T.: Comparison of the reliability of scoring methods of multiple-response items, matching items, and sequencing items. CADMO. (2012).