Usage of Multiple-Choice Items in Summative Examinations: Questionnaire Survey Among German Undergraduate Dental Training Programmes

Lena Rössler, Manfred Herrmann, Annette Wiegand, Philipp Kanzow

Submitted to: JMIR Medical Education
on: March 07, 2024

Table of Contents

Original Manuscript **5**
Supplementary Files **23**
Figures **24**
Figure 1 ..... **25**
Multimedia Appendixes **26**
Multimedia Appendix 1 ..... **27**
TOC/Feature image for homepages **28**
TOC/Feature image for homepage 0 ..... **29**

Usage of Multiple-Choice Items in Summative Examinations: Questionnaire Survey Among German Undergraduate Dental Training Programmes

Lena Rössler[1]; Manfred Herrmann[2] Dr rer nat; Annette Wiegand[1] Pro f Dr med dent; Philipp Kanzow[1] MSc, Dr rer medic, PD Dr med dent

[1]Department of Preventive Dentistry, Periodontology and Cariology University Medical Center Göttingen Göttingen DE
[2]Study Deanery University Medical Center Göttingen Göttingen DE

Corresponding Author:
Philipp Kanzow MSc, Dr rer medic, PD Dr med dent
Department of Preventive Dentistry, Periodontology and Cariology
University Medical Center Göttingen
Robert-Koch-Str 40
Göttingen
DE

Abstract

Background: Multiple-choice examinations are frequently employed among German dental schools. However, details regarding the used item types and applied scoring methods are lacking.

Objective: We aimed to gain an insight into the current usage of multiple-choice items (ie, questions) in summative examinations in German undergraduate dental training programmes.

Methods: A paper-based 10-item questionnaire regarding the employed assessment methods, multiple-choice item types, and applied scoring methods was designed. The pilot-tested questionnaire was mailed to the Deans of Studies and to the Heads of Department of Operative/Restorative Dentistry at all 30 dental schools in Germany in February 2023. Statistical analysis was performed using Fisher's exact test (P<.05).

Results: The response rate amounted to 90.0% (27/30 dental schools). All respondent dental schools employed multiple-choice examinations for summative assessments. Examinations were delivered electronically by 70.4% (19/27) of the dental schools. Almost all dental schools used single-choice Type A items (88.9%) which accounted for the largest number of items in about half of the dental schools. Further item types (eg, conventional multiple-select items, Multiple-True-False, Pick-N) were only used by fewer dental schools (?66.7%, up to 18 out of 27 dental schools). For the multiple-select item types, the applied scoring methods varied considerably (ie, awarding [intermediate] partial credit, requirements for partial credit). Dental schools with the possibility of electronic examinations used multiple-select items slightly more often (73.7%, 14/19 vs. 50.0%, 4/8). However, this difference was statistically not significant (P=.375). Dental schools used items either individually or as key feature problems consisting of a clinical case scenario followed by a number of items focusing on critical treatment steps (55.6%, 15/27). Not a single school employed alternative testing methods (eg, answer-until-correct). A formal item review process was established at about half of the dental schools (55.6%, 15/27).

Conclusions: Summative assessment methods among German dental schools vary widely. Especially, a large variability regarding the use and scoring of multiple-select multiple-choice items was found.

(JMIR Preprints 07/03/2024:58126)
DOI: https://doi.org/10.2196/preprints.58126

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
Please make my preprint PDF available to anyone at any time (recommended).
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

✓ **No, I do not wish to publish my submitted manuscript as a preprint.**

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Paper

Usage of Multiple-Choice Items in Summative Examinations: Questionnaire Survey Among German Undergraduate Dental Training Programmes

Abstract

Background: Multiple-choice examinations are frequently employed among German dental schools. However, details regarding the used item types and applied scoring methods are lacking.

Objective: We aimed to gain an insight into the current usage of multiple-choice items (ie, questions) in summative examinations in German undergraduate dental training programmes.

Methods: A paper-based 10-item questionnaire regarding the employed assessment methods, multiple-choice item types, and applied scoring methods was designed. The pilot-tested questionnaire was mailed to the Deans of Studies and to the Heads of Department of Operative/Restorative Dentistry at all 30 dental schools in Germany in February 2023. Statistical analysis was performed using Fisher's exact test (P<.05).

Results: The response rate amounted to 90.0% (27/30 dental schools). All respondent dental schools employed multiple-choice examinations for summative assessments. Examinations were delivered electronically by 70.4% (19/27) of the dental schools. Almost all dental schools used single-choice Type A items (88.9%) which accounted for the largest number of items in about half of the dental schools. Further item types (eg, conventional multiple-select items, Multiple-True-False, Pick-N) were only used by fewer dental schools (≤66.7%, up to 18 out of 27 dental schools). For the multiple-select item types, the applied scoring methods varied considerably (ie, awarding [intermediate] partial credit, requirements for partial credit). Dental schools with the possibility of electronic examinations used multiple-select items slightly more often (73.7%, 14/19 vs. 50.0%, 4/8). However, this difference was statistically not significant (P=.375). Dental schools used items either individually or as key feature problems consisting of a clinical case scenario followed by a number of items focusing on critical treatment steps (55.6%, 15/27). Not a single school employed alternative testing methods (eg, answer-until-correct). A formal item review process was established

at about half of the dental schools (55.6%, 15/27).

Conclusions: Summative assessment methods among German dental schools vary widely. Especially, a large variability regarding the use and scoring of multiple-select multiple-choice items was found.

Keywords: alternate-choice; assessment; best-answer; dental; dental schools; dental training; education; educational assessment; educational measurement; examination; German; Germany; k of n; Kprim; K'; medical education; medical student; MTF; Multiple-True-False; multiple choice; multiple-select; Pick-N; scoring; scoring system; single choice; single response; test; testing; true/false; true-false; Type A; Type K; Type K'; Type R; Type X; undergraduate; undergraduate curriculum; undergraduate education

Introduction

Summative examinations of theoretical knowledge are an integral part of university degree programmes. As they are intended to assess examinees' ability regarding predefined learning objectives, they should reflect examinees' true knowledge as closely as possible. To assess examinees objectively and efficiently, multiple-choice examinations were described as early as 1916 [1,2]. To date, these types of examinations have been expanded by further item types, and multiple-choice examinations are frequently used within higher education including but not limited to dental training programmes [3-5]. Multiple-choice items (ie, questions) can be subdivided into single-choice items (eg, Type A, Type K, Type R, alternate-choice), and multiple-select items (eg, Pick-N, Multiple-True-False [Type K']) [6]. While dichotomous scoring (ie, 1 full credit point is awarded if examinees mark the correct answer option or statements, otherwise no credit is awarded) is most commonly proposed for single-choice items [7], scoring methods for multiple-select items are more heterogeneous: Besides dichotomous scoring, further scoring methods resulting in (intermediate) partial credit or even negative points (ie, malus points) have been described [8,9].

Besides paper-based examinations, examinations are nowadays frequently delivered electronically. While electronic examinations are well perceived by examinees [10], comprehensive studies regarding their effectiveness are still lacking [11]. However, the use of different examination software (eg, UCAN's CAMPUS examination software) might improve the ease of multiple-choice examinations, accelerate the evaluation of examinations and item analysis, and allow for more complex scoring algorithms. Despite the benefits associated with electronic examinations, the availability of hardware and software at the level of individual institutions might limit its use.

In Germany, the revised undergraduate dental curriculum consists of ten semesters and includes pre-clinical training (4 semesters), training using simulators or phantom heads (2 semesters), and clinical training (4 semesters). Following the state examinations after each part (ie, at after the fourth, sixth, and tenth semester), students receive their license ("Approbation") to practice dentistry. Besides practical skills, theoretical knowledge is taught within the undergraduate dental curriculum,

and students' ability is often assessed using written multiple-choice examinations. However, such examinations are not standardized among German dental schools. While general recommendations exist for their design and evaluation [12,13], details such as suitable item types and applied scoring methods are often defined in local examination guidelines at the level of individual dental schools. However, these details might impact examinees' scoring results [5]. To the best of our knowledge, a comprehensive overview regarding the used item types and applied scoring methods at German dental schools does not exist.

Therefore, this study aimed to gain an insight into the current usage of multiple-choice items in summative examinations in German undergraduate dental training programmes. The null-hypothesis is that the use of digital examinations does not impact the use of more-complex (ie, multiple-select) multiple-choice items.

Methods

Ethical Considerations

The study was designed as a prospective investigation. In preparation for the investigation, the websites of all German dental schools were screened (n=30) and the names of the Heads of Department of Operative/Restorative Dentistry and the Deans of Studies were noted for later procedure.

The study was performed after approval by the local ethics committee of the University Medical Center Göttingen (no. 22/1/23). The participation in this study was voluntary, and participants gave their informed consent for the anonymous evaluation of the provided answers by returning the questionnaires. Participants did not receive any incentives or compensation.

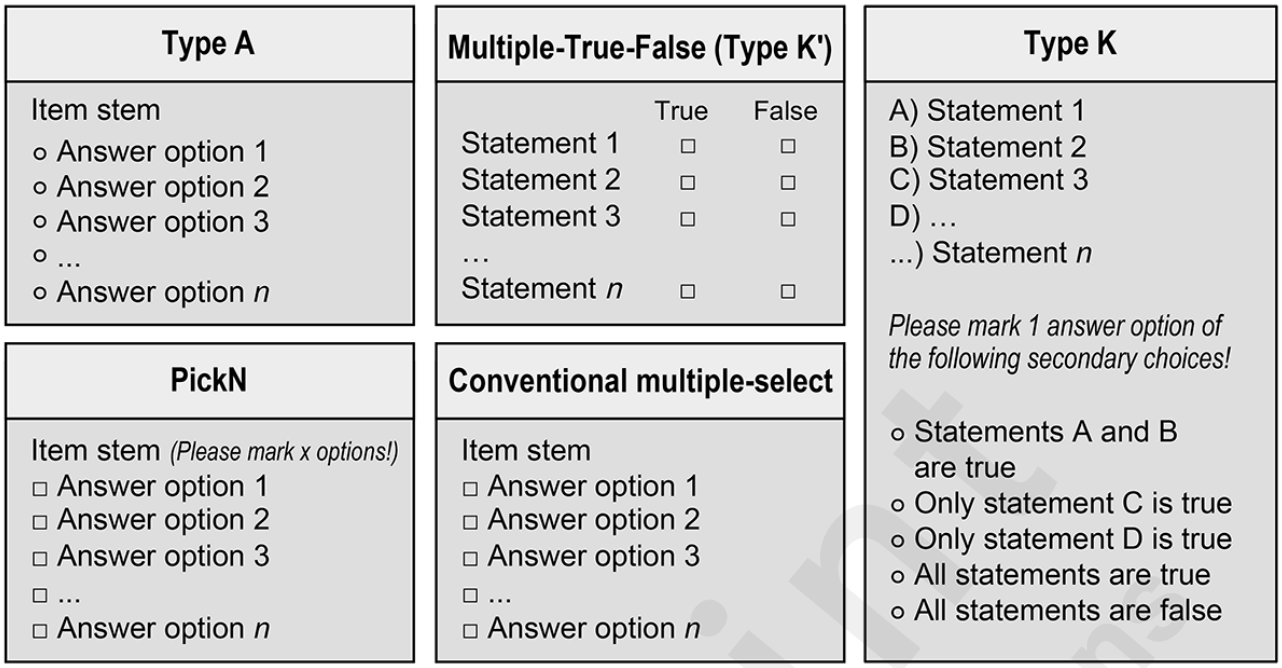
Questionnaire

A paper-based questionnaire, consisting of 10 items about the construction and evaluation of

summative examinations was jointly designed by the authors and pilot-tested in the University Medical Center Göttingen (Multimedia Appendix 1). Both closed and open-ended items were used. The opening questions related to different examination types used for the summative assessment of theoretical knowledge, and whether or not electronic examinations were being used. Additionally, it was asked whether the examination items undergo a formal review process and if so, the participants had the chance to give a brief description of this procedure. The more specific questions related to the types of multiple-choice items utilized and asked for the relative percentage to which these items were being used. Furthermore, the participants were asked to describe the applied scoring methods for each of the item types used. Finally, participants were provided with a text field open for comments and/or their contact details (ie, if required for further clarification) and were asked to supply a copy of their local examination guideline or programme regulations.

Following the evaluation of the pilot-survey among 5 dentists at the University Medical Center Göttingen, the questionnaire was slightly modified for clarification, printed, and mailed to (1) the Heads of Department of Operative/Restorative Dentistry and to (2) the Deans of Studies on February 1st, 2023. The wording was slightly adjusted for each recipient: (1) "used in your department" versus (2) "permitted at your dental school". Mailings included a personalized cover letter, an overview illustrating different multiple-choice item types (Figure 1), and a stamped return envelope. The survey was closed after 12 weeks. Non-responders were reminded once 6 weeks after the initial distribution of the questionnaires.

Figure 1. Exemplary presentation of the most commonly used multiple-choice item types referenced in the questionnaire. Round marking boxes represent 1 answer option to be selected (1 out of X), while square marking boxes imply that multiple answer options or statements (x out of X) can be chosen.



Statistical Analysis

First, data was manually transferred into a digital chart using a piloted spreadsheet containing columns for each item of the questionnaire. This step was independently performed by 2 authors (LR, PK). In case of disagreement, data was repeatedly extracted from the returned questionnaires.

In case of disagreement between the Heads of Department of Operative/Restorative Dentistry and the Deans of Studies, results were based on the responses from the Heads of Department. For further clarification, responses were cross-validated with the supplied / publicly available examination guidelines and programme regulations. If required, respondents were contacted for further clarification if they had agreed to do so previously.

Second, statistical analysis was performed using the software SPSS Statistics (Macintosh version 29.0.0.0; IBM Corp). The effect of delivering digital examinations on the use of multiple-select items was assessed using Fisher’s exact test. The level of significance was set at 0.05.

Results

In total, responses from 27 dental schools were received yielding a response rate of 90% (27/30). More specifically, 25 Departments of Operative/Restorative Dentistry and 17 Deans of Studies replied. All dental schools responded that they use written multiple-choice examinations for the assessment of examinees' theoretical knowledge. Therefore, subsequent results are based on the number of respondent dental schools.

Multiple-choice Items Used

The most commonly used multiple-choice item types at German dental schools were single-choice Type A or Type $A_{negative}$ items with 3 to 6 answer options (88.9%, 24/27). Pick-N items (ie, the number of answer options to be selected is known to examinees) were reported to contain between 3 and 26 answer options and were used by 66.7% (18/27) of dental schools. Type K items were reported to contain between 3 and 6 statements and were used by 51.9% (14/27) of the dental schools. Multiple-True-False (also known under further names such as Kprim, Type K', or Type X) and conventional multiple-select items (ie, the number of answer options to be selected is unknown to examinees) were reported to contain between 4 and 6 statements / answer options and were both used by 44.4% (12/27) of the dental schools. The use of further item types is shown in Table 1.

Table 1. Different multiple-choice item types for the assessment of theoretical knowledge at different German dental schools.

Item Type	Dental schools, n (%)
Type A	24 (88.9)
Pick-N	18 (66.7)
Type K	14 (51.9)

Conventional multiple-select	12 (44.4)
Multiple-True-False (Type K')	12 (44.4)
Type R	6 (22.2)
Alternate-choice	4 (14.8)

Examination Setting

Key feature problems consisting of a clinical case scenario followed by a number of items focusing on critical treatment steps were used by about half of the dental schools (55.6%, 15/27). Not a single school employed alternative testing methods (eg, answer-until-correct). Also, a formal item review process prior to the delivery of the examination was only established at about half of the dental schools (55.6%, 15/27).

Delivery of Examinations

The percentage of dental schools which deliver examinations electronically amounted to 70.4% (19/27). However, the software used by the dental schools differed: A dedicated examination software (ie, UCAN's CAMPUS or tEXAM, Q-Exam) was used by 8 dental schools, while learning management systems such as Moodle, ILIAS, or OpenOLAT were used by 7 dental schools for the purpose of examination delivery. The remaining 4 dental schools did not provide any information regarding the utilized examination software.

Dental schools with the possibility of electronic examinations used multiple-select items slightly more often (73.7%, 14/19 vs. 50.0%, 4/8). However, this difference was statistically not significant (*P*=.375).

Applied Scoring Methods

All dental schools scored single-choice items (ie, Type A, Type $A_{negative}$, Type K, Type R, alternate-choice) dichotomously (ie, 1 full credit point is awarded if examinees mark the correct answer option or statements, otherwise no credit is awarded).

Scoring of multiple-select items was more heterogeneous and no single scoring method which was commonly used was identified: Some dental schools utilized scoring algorithms resulting in partial (ie, 0.5 credit points) or intermediate partial credit (ie, 1/n partial credit for each correct response) besides dichotomous scoring on multiple-select items. However, scoring methods resulting in negative points (ie, malus points) were not used at any location.

Discussion

Principal Findings

The aim of this study was to gain an insight into summative assessment methods using multiple-choice items at German dental schools. The purpose of summative assessment is to evaluate examinees' knowledge at the end of a course by comparing their scores to a predefined standard (ie, cut-off score) [14]. Our results demonstrate that all respondent dental schools employ multiple-choice examinations for summative assessment of theoretical knowledge. Besides individual items, about half of the dental schools also employ key feature problems.

Single-choice Type A items are the most popular item type used at German dental schools. These items are employed by almost every respondent dental school and often account for the largest number of items at the respective dental schools. This might be explained by the demand for ease of scoring (ie, dichotomous scoring, no partially correct responses).

Multiple-select item types such as Pick-N or Multiple-True-False are employed by fewer dental schools. For these item types, the applied scoring methods vary considerably: Some dental schools award partial or even intermediate partial credit for partially correct responses while others do not.

However, the exact cut-off levels and scoring methods for partial credit differed. For example, Partial Scoring 50% (PS_{50}) was used by some dental schools for Pick-N items: In these cases, 1 full credit point is awarded if all answer options are marked correctly, 0.5 credit points are awarded if at least half of the true answers option are marked, otherwise no credit is awarded [9,15]. Also, a similar scoring method named Half-point Scoring was used by some dental schools for Multiple-True-False and conventional multiple-select items: 1 full credit point is awarded if all statements / answer options are marked correctly, 0.5 credit points are awarded if the response to 1 statement / answer option is incorrect, otherwise no credit is awarded [8,16]. In addition, some dental schools awarded intermediate partial credit on multiple-select items: In case of Partial Scoring 1/n ($PS_{1/n}$), 1/n partial credit was awarded for each correct response [8,9]. Some dental schools also subtracted 1/n partial credit for each incorrect response (Blasberg-Method) [8,9,17].

As a result, scoring of multiple-select items at different German dental schools can be considered very heterogeneous. This is not surprising, as a fast number of different scoring methods for multiple-select items has been described in literature [8,9]. As stated previously, it is not possible to suggest a single versatile scoring method. Different requirements as defined in dental schools' local examination guidelines (eg, fixed pass-mark, fixed proportion of true answer options) impact the scoring method to be selected. Regarding jurisdictional requirements, scoring methods resulting in negative points (ie, malus points) must not be used in Germany [13]. Consequently, not a single dental school uses scoring methods resulting in malus points. However, almost half of the dental schools do not employ a formal item review process. A formal review processes is recommended prior to the delivery of the examinations and might further improve the quality and overall validity of the examinations.

In addition, 70.4% (19/27) of all dental schools stated to deliver examinations electronically. While the electronic delivery of examinations allows for automatic scoring and more complex scoring methods (ie, within the context of multiple-select items), no statistically significant relation between the type of delivery (paper-based vs. electronic) and the use of multiple-select item types was

found. Thereby, our results fail to reject the null-hypothesis. This might be explained by the software used for the delivery and scoring of electronic examinations: It was found that dental schools employ learning managements systems such as Moodle, ILIAS, or OpenOLAT besides dedicated examination software such as UCAN's CAMPUS, tEXAM, or Q-Exam for the delivery and scoring of summative assessments. This is of relevance, as learning managements systems usually support fewer item types and scoring methods than dedicated examination software [8,9]. As a results, electronic delivery of examinations does not necessarily result in an increased use of multiple-select items.

Interestingly, not a single dental school employed alternative testing methods which deviate from the standard setting during examinations (ie, examinees marking the answer options or statements they believe to be correct or true but receive no immediate feedback regarding correctly or incorrectly marked answer options or statements). Within multiple-choice examinations, alternative testing methods such as confidence weighting scoring (ie, examinees are requested to indicate the degree of confidence in their marking) [18], elimination scoring (ie, examinees are instructed to mark the incorrect instead of correct answer options) [19], or answer-until-correct [20,21] have been described in the literature. Within the answer-until-correct method, examinees receive immediate feedback and examinees may correct their marking on previously incorrectly marked items, thereby still receiving partial credit. However, the benefit of such testing methods within the field of dental education is questionable. Dental schools' examinees are becoming future dentists. While treating patients, dentists are required to make informed choices and dentists might not always have a second chance without potentially harming their patients. Also, such alternative testing methods benefit from the electronic delivery of examinations and set even higher requirements towards the used examination software.

Strengths and Limitations

To the best of our knowledge, this is the first study to systematically assess the use and scoring of

multiple-choice item types in summative examinations among German dental schools. A number of strengths are present: First, a pre-tested questionnaire was used. Second, our questionnaire survey study yielded a high response rate of 90% (27/30). Third, our results might be considered representative for current use of multiple-choice items in summative examinations among German dental schools.

Nevertheless, limitations are also present: First, our questionnaire focused on multiple-choice items, therefore the usage of other assessment types (eg, objective structured clinical examinations, oral examinations) remains unknown. Second, the number of dental schools in Germany is limited. Thereby, results from Fisher's exact test might be underpowered despite the high response rate. Also, this study could not control for potential confounders (eg, location, number of students per dental school) due the overall low number of dental schools. Third, transferability and generalizability to other educational settings might be limited due to different jurisdictional requirements or overall lower importance of written examinations.

Future Directions

Since 2021, new dental licensing regulations ("Approbationsordnung") are in effect which restructured the undergraduate dental curriculum in Germany. For the first time, a nationwide written board examination with single-choice items takes place at the end of all undergraduate dental programmes (ie, after the tenth semester) [22]. Therefore, multiple-choice examinations in general and especially single-choice Type A items will remain a popular format for summative examinations among German undergraduate dental programmes. Ideally, examinees already become familiar with single-choice Type A items during their studies. Therefore, all dental schools should employ single-choice Type A items to adequately prepare their students for the final board examination.

Nevertheless, additional examinations (eg, objective structured clinical or practical examinations) are required to test examinees' practical skills [3]. Regardless of the employed item type, multiple-choice examinations are not suitable to assess the higher levels Miller's Pyramid of clinical

competence (ie, does, shows how) [23].

Conclusion

While students from almost all dental schools can be expected to be familiar with single-choice Type A items, summative assessment techniques of theoretical knowledge differ widely among German dental schools. Especially, a large variability regarding the use and scoring of multiple-select multiple-choice items was found. Also, implementing a formal item review process might further improve the quality and overall validity of the examinations.

Acknowledgments

The authors acknowledge support by the Open Access Publication Funds of Göttingen University. The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. Generative artificial intelligence was not used in any portion of this manuscript.

Data Availability

All data generated during or analyzed during this study are included in this published article.

Authors’ Contributions

LR, MH, AW, and PK contributed to the study’s conception and designed the questionnaire. LR and PK transferred the data. PK performed statistical analyses. All authors interpreted the data, critically revised the manuscript, and approved the final version of the manuscript.

Conflicts of Interest

PK is Associate Editor of JMIR Medical Education at the time of this publication. Other authors have no competing interests to declare.

Abbreviations

$PS_{1/n}$: Partial Scoring 1/n.

PS_{50}: Partial Scoring 50%.

Multimedia Appendix 1: Authors' translation of the utilized questionnaire which was originally distributed in German.

References

1. Kelly FJ. The Kansas silent reading test. J Educ Psychol 1916; 7(2):63-80. doi:10.1037/h0073542

2. Ruch GM, Stoddard GD. Comparative reliabilities of five types of objective examinations. J Educ Psychol 1925; 16(2):89-103. doi:10.1037/h0072894

3. Gerhard-Szep S, Güntsch A, Pospiech P, Söhnel A, Scheutzel P, Wassmann T, Zahn T. Assessment formats in dental medicine: an overview. GMS J Med Educ 2016; 33(4):Doc65. PMID:27579365 doi:10.3205/zma001064

4. Kanzow P, Schuelper N, Witt D, Wassmann T, Sennhenn-Kirchner S, Wiegand A, Raupach T. Effect of different scoring approaches upon credit assignment when using Multiple True-False items in dental undergraduate examinations. Eur J Dent Educ 2018; 22(4):e669-e678. PMID:29934980 doi:10.1111/eje.12372

5. Kanzow P, Schmidt D, Herrmann M, Wassmann T, Wiegand A, Raupach T. Use of multiple-select multiple-choice items in a dental undergraduate curriculum: retrospective study involving the application of different scoring methods. JMIR Med Educ 2023; 9:e43792. PMID:36841970 doi:10.2196/43792

6. Krebs R. Prüfen mit Multiple Choice: Kompetent planen, entwickeln, durchführen und auswerten [Testing with multiple choice: plan, develop, implement, and evaluate competently]. Bern, Switzerland: Hogrefe; 2019. ISBN:9783456859026

7. Kanzow AF, Schmidt D, Kanzow P. Scoring single-response multiple-choice items: scoping review and comparison of different scoring methods. JMIR Med Educ 2023; 9:e44084. PMID:37001510 doi:10.2196/44084

8. Schmidt D, Raupach T, Wiegand A, Herrmann M, Kanzow P. Relation between examinees' true knowledge and examination scores: systematic review and exemplary calculations on Multiple-True-False items. Educ Res Rev 2021; 34:100409. doi:10.1016/j.edurev.2021.100409

9. Schmidt D, Raupach T, Wiegand A, Herrmann M, Kanzow P. Relation between examinees' true knowledge and examination scores: systematic review and exemplary calculations on Pick-N

items. Educ Res Rev 2022; 37:100483. doi:10.1016/j.edurev.2022.100483

10. Nardi A, Ranieri M. Comparing paper-based and electronic multiple-choice examinations with personal devices: impact on students' performance, self-efficacy and satisfaction. Brit J Educ Technol 2018; 50(3):1495-1506. doi:10.1111/bjet.12644

11. Way WD, Davis LL, Keng L, Strain-Seymour E. Increasing the accessibility of assessments through technology. In: Drasgow F, editor. Technology and testing: improving educational and psychological measurement. New York: Routledge; 2015. p. 217-234. ISBN:9781315871493

12. Jünger J, Just I. Recommendations of the German Society for Medical Education and the German Association of Medical Faculties regarding university-specific assessments during the study of human, dental and veterinary medicine. GMS Z Med Ausbild 2014; 31(3):Doc34. PMID:25228936 doi:10.3205/zma000926

13. Kubinger KD. Gutachten zur Erstellung „gerichtsfester" Multiple-Choice-Prüfungsaufgaben [Expert opinion on the creation of "lawful" multiple-choice items]. Psychol Rundschau 2014; 65(3):169-178. doi:10.1026/0033-3042/a000218

14. Scriven M. The methodology of evaluation. In: Tyler RW, Gagné RM, Scriven M, editors. Perspectives of curriculum evaluation. Chicago: Rand McNally; 1967. ISBN:9780528618109

15. Bauer D, Holzer M, Kopp V, Fischer MR. Pick-N multiple choice-exams: a comparison of scoring algorithms. Adv Health Sci Educ Theory Pract 2011; 16(2):211-221. PMID:21038082 doi:10.1007/s10459-010-9256-1

16. Vorkauf H. Teilpunktbewertung bei K'-Items [Partial credit scoring of Multiple-True-False items]. Jahresbericht 1986 der Gruppe Medizinalprüfungen und der Gruppe Statistik und EDV. Bern, Switzerland: Institut für Ausbildungs- und Examensforschung, Medizinische Fakultät der Universität Bern; 1987. p. 44-48.

17. Blasberg R, Güngerich U, Müller-Esterl W, Neumann D, Schappel S. Erfahrungen mit dem Fragentyp „k aus n" in Multiple-Choice-Klausuren [Experiences with item type "k from n" in multiple-choice tests]. Med Ausbild 2001; 18(S1):73-76.

18. Hevner K. A method of correcting for guessing in true-false tests and empirical evidence in

support of it. J Soc Psychol 1932; 3(3):359-362. doi:10.1080/00224545.1932.9919159

19. Collet LS. Elimination scoring: an empirical evaluation. J Educ Meas 1971; 8(3):209-214. doi:10.1111/j.1745-3984.1971.tb00927.x

20. Gilman DA, Ferry P. Increasing test reliability through self-scoring procedures. J Educ Meas 1972; 9(3):205-207. doi:10.1111/j.1745-3984.1972.tb00953.x

21. Hanna GS. Incremental reliability and validity of multiple-choice tests with an answer-until-correct procedure. J Educ Meas 1975; 12(3):175-178. doi:10.1111/j.1745-3984.1975.tb01019.x

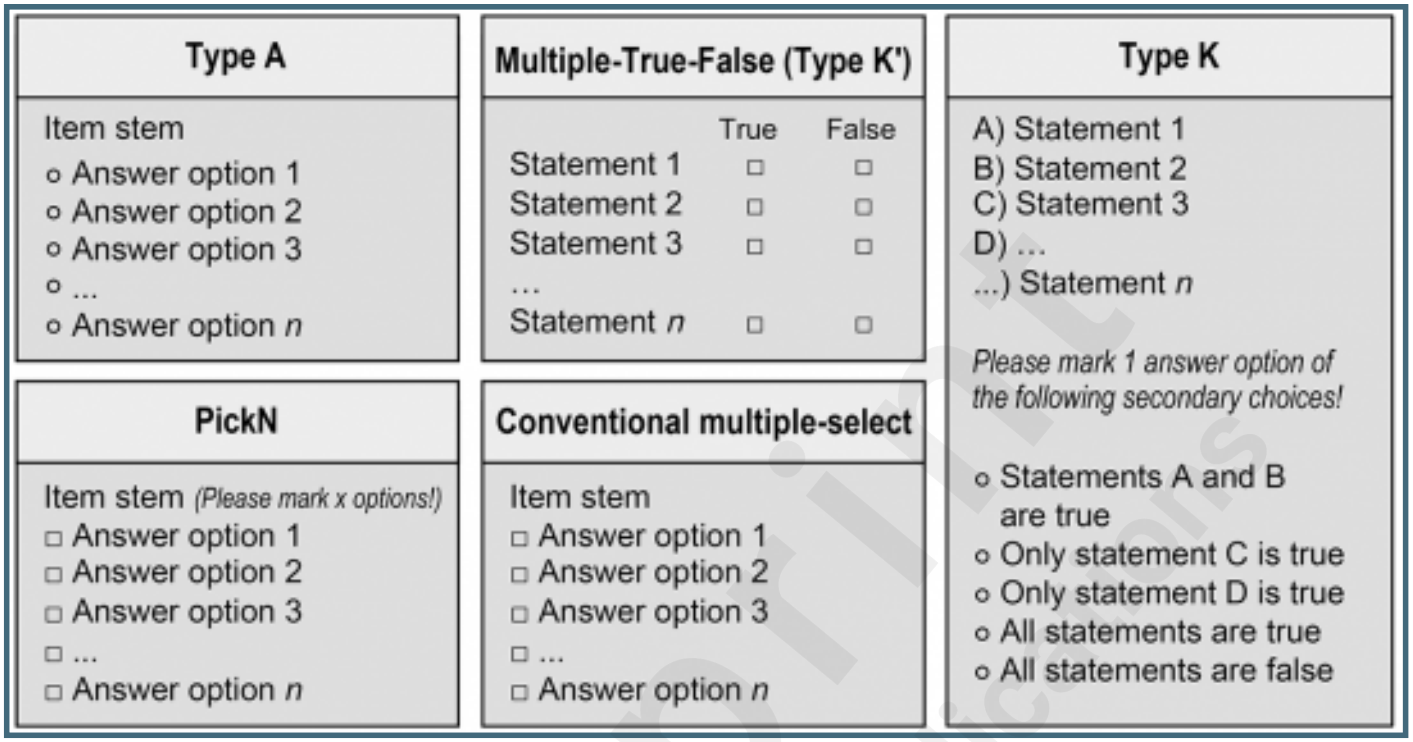
22. Schmitz UJ, Daubländer M, editors. Übertragung papierbasierter Multiple Choice Aufgaben in die digitale Welt – ein Weg zur Verbesserung der Prüfungsqualität? [Transferring paper-based multiple-choice items to the digital world – a way to improve examination quality?]. Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA) 2022; Halle (Saale), Germany.

23. Miller GE. The assessment of clinical skills/competence/performance. Acad Med 1990; 65(9):S63-S67. PMID:2400509 doi:10.1097/00001888-199009000-00045

Supplementary Files

Figures

Exemplary presentation of the most commonly used multiple-choice item types referenced in the questionnaire. Round marking boxes represent 1 answer option to be selected (1 out of X), while square marking boxes imply that multiple answer options or statements (x out of X) can be chosen.



Multimedia Appendixes

Authors' translation of the utilized questionnaire which was originally distributed in German.
URL: http://asset.jmir.pub/assets/0a49ceec6fe0288583362e0c10976ddc.pdf

TOC/Feature image for homepages

Electronic multiple-choice examination at the University Medical Center Göttingen.