

# Digital transformation of the Turkish national neurology board examination: Implementation and candidates' feedback

S. Ayhan Çalışkan<sup>1,2</sup>, Gülşen Taşdelen Teker<sup>3</sup>, Hatice Mavioglu<sup>4</sup>, Özgül Ekmekci<sup>5</sup>,  
Figen Gökçay<sup>5</sup>, Figen Eşmeli<sup>6</sup>, Semiha Kurt<sup>7</sup>, Füsun Ferda Erdoğan<sup>8</sup>, Meltem Demirkiran<sup>9</sup>,  
Bijen Nazlıel<sup>10</sup>, Esen Saka<sup>11</sup>

<sup>1</sup>Department of Medical Education, United Arab Emirates University, College of Medicine and Health Sciences, Al Ain, United Arab Emirates

<sup>2</sup>Department of Medical Education, İzmir University of Economics, Faculty of Medicine, İzmir, Türkiye

<sup>3</sup>Department of Medical Education and Informatics, Hacettepe University Faculty of Medicine, Ankara, Türkiye

<sup>4</sup>Department of Neurology, Celal Bayar University Faculty of Medicine, Manisa, Türkiye

<sup>5</sup>Department of Neurology, Ege University Faculty of Medicine, İzmir, Türkiye

<sup>6</sup>Department of Neurology, Balıkesir University Faculty of Medicine, Balıkesir, Türkiye

<sup>7</sup>Department of Neurology, Tokat Gaziosmanpaşa University Faculty of Medicine, Tokat, Türkiye

<sup>8</sup>Department of Neurology, Erciyes University Faculty of Medicine, Kayseri, Türkiye

<sup>9</sup>Department of Neurology, Çukurova University Faculty of Medicine, Adana, Türkiye

<sup>10</sup>Department of Neurology, Gazi University Faculty of Medicine, Ankara, Türkiye

<sup>11</sup>Department of Neurology, Hacettepe University Faculty of Medicine, Ankara, Türkiye

## ABSTRACT

**Objectives:** This study aimed to present the implementation of the computer-based Turkish National Neurology Board Examination (TNNBE) process, which was digitalized by the Turkish Board of Neurology (TBN) using an open-source learning management system to improve accessibility, and the feedback from candidates.

**Materials and methods:** Neurology academics submitted items to the exam pool. From this pool, TBN members selected 79 multiple-choice questions (MCQs) and 10 key feature problems (KFPs), each containing two to four items (totaling 30 questions), through 20 h of online meetings and discussions to ensure content validity. Standard setting was applied using the Nedelsky (MCQ) and Angoff (KFP) methods. Instructions and a sample exam were created and sent to candidates for piloting purposes. Sixty neurologists (35 females, 25 males; mean age: 30.6±1.48 years; range, 28 to 36 years) who fulfilled the eligibility criteria participated in the exam, which was conducted in December 2023 at one venue, under supervision, with online security measures such as individual passwords for login, question shuffling, and browser lockdown. The total exam time was 130 min, divided into MCQs followed by KFPs. Each question was worth 1 point, with a maximum of 100 points obtainable. The MCQ scores were calculated automatically by the learning management system, while KFP responses were downloaded and scored by two different board members, with a final consensus mark determined through discussion among all board members. The final score was the sum of MCQ and KFP scores. Candidates' feedback was obtained via an online survey using a 9-point scale (1=strongly disagree/very bad; 9=strongly agree/very good).

**Results:** Seven MCQs were omitted from the exam set due to various reasons. The mean scores were 47.73±6.61 for MCQs, 16.09±3.82 for KFPs, and 63.83±8.79 overall. Thirty (50.0%) candidates scored higher than the minimum acceptable level of performance (65/100) and passed the exam. The mean score percentage for the MCQ section (68.2%) was significantly higher than for the KFP section (53.1%;  $p<0.001$ ). Cronbach's alpha coefficients for the MCQ (63 items) and KFP (30 items) sections were 0.75 and 0.67, respectively. Candidates provided positive feedback ( $n=56$ ), indicating that the exam venue was comfortable ( $\bar{X}=6.02\pm2.42$ ), the digital format was easy to use ( $\bar{X}=5.44\pm2.62$ ), and the exam user interface was convenient ( $\bar{X}=5.96\pm2.45$ ). The highest satisfaction was for the inclusion of clinical case questions ( $\bar{X}=6.63\pm2.21$ ). Candidates also found the KFP section ( $\bar{X}=6.59\pm1.80$ ) more challenging than the MCQ section ( $\bar{X}=6.13\pm1.61$ ).

**Conclusion:** The computer-based TNNBE effectively streamlined the exam process and received positive feedback from candidates, particularly for its user interface and inclusion of KFPs. However, KFP scoring was still challenging due to the range and small differences in formulating the acceptable answers, which made it difficult to standardize scoring by machines. In addition, implementing computer-based exams, particularly in large-scale settings, require advanced technology and logistical planning, which can be costly and challenging to manage.

**Keywords:** Board exam, computer-based exam, key feature problems, multiple-choice questions.

**Correspondence:** S. Ayhan Çalışkan, MD, PhD. Department of Medical Education, United Arab Emirates University, College of Medicine and Health Sciences, 15551 Al Ain, United Arab Emirates.

E-mail: ayhanca@gmail.com

**Received:** December 23, 2024 **Accepted:** April 07, 2025 **Published online:** May 05, 2025

**Cite this article as:** Çalışkan SA, Taşdelen Teker G, Mavioglu H, Ekmekci Ö, Gökçay F, Eşmeli F, et al. Digital transformation of the Turkish national neurology board examination: Implementation and candidates' feedback. Turk J Neurol 2025;31(3):270-277. doi: 10.55697/tnd.2025.383.

The Turkish National Neurology Board Examination (TNNBE) has been conducted since 2004 by the Exam Commission of the Turkish Board of Neurology (TBN) within the Turkish Neurological Society.<sup>[1,2]</sup> Until 2015, the exam relied exclusively on multiple-choice questions (MCQs), which are highly effective for assessing broad knowledge areas, as they allow the test to cover a wide range of content and enhance content-related validity.<sup>[3]</sup> This approach supports making valid inferences about the entire content domain. Additionally, MCQs are widely used due to their high reliability and ease of grading, offering accuracy, consistency, and speed.<sup>[4]</sup> However, poorly crafted or flawed MCQs may test trivial content at a low cognitive level.<sup>[5,6]</sup> Therefore, the types of questions included in specific assessments should be selected based on their distinct advantages and limitations. An effective assessment process employs a variety of methods, each carefully designed to meet the evaluation's specific requirements.<sup>[7]</sup>

As a result, starting in 2015, items in the key feature problems (KFPs) format were introduced to the TNNBE to better assess candidates' clinical decision-making skills.<sup>[8]</sup> Key feature problems are context-rich questions that require candidates to integrate multiple pieces of information to arrive at clinically meaningful decisions.<sup>[7]</sup> They are particularly valuable for evaluating clinical decision-making as they assess not only medical knowledge but also the ability to apply it in clinical contexts. This involves making critical decisions at specific points during the evaluation and management of a case. These critical points define the "key features" of the problem.<sup>[9]</sup>

The key features test format, initially introduced at the Cambridge Conference in 1984, was incorporated into the Medical Council of Canada Qualifying Examination Part I in 1992. This addition aimed to replace the PMPs (patient management problems) and reduce exclusive reliance on MCQs for assessing licensure qualifications.<sup>[10-12]</sup> For a similar purpose, KFPs were added to the TNNBE.

The most recent change to the exam occurred in 2023, transitioning it from a paper-and-pencil format to a computer-based test. During the coronavirus disease 2019 (COVID-19) pandemic, academic institutions rapidly transitioned educational activities, including exams, to an electronic learning format. In subsequent periods, institutions voluntarily continued online educational activities due to their advantages in preparation, administration, and scoring.

This study aimed to present the implementation of the computer-based TNNBE, share the exam results, and highlight feedback received from candidates regarding the computer-based format.

## MATERIALS AND METHODS

The cross-sectional study was conducted with individuals who completed neurology residency training at authorized educational institutions and could provide certification confirming their eligibility to apply for the TNNBE. Additionally, neurology residents in their final year of training could apply for the board exam by submitting their residency logbook, including a list of minimum required practical procedures, approved by supervising neurologists, along with documents verifying the completion of necessary rotations, to the examination committee. The committee reviewed these documents, and those deemed eligible were granted permission to take the exam. Of the 96 applicants, 60 eligible neurologists (35 females, 25 males; mean age: 30.6±1.5 years; range, 28 to 36 years) participated in the exam held on December 13, 2023.

Ethical approval was obtained from the Ege University Scientific Research Ethics Boards (date: 14.12.2023, no: 23-12T/4). We confirm that all methods used in this study were carried out in accordance with relevant guidelines and regulations.

Of the participants, 48 (60%) were in their final year of residency, while 12 (20%) were neurology specialists who had already completed their residency. Among these specialists, seven had completed their residency in the same year as the exam (2023), four had finished one year earlier, and one had completed two years prior.

An analysis of the institutions where participants completed their residency revealed two distinct types of training institutions in Türkiye: university hospitals and education and research hospitals. Significant differences exist between these institution types. Education and research hospitals generally handle a broader range and higher volume of patients compared to university hospitals. Regarding the institutions where participants completed their residency, 19 participants were either currently training or had completed their residency at an education and research hospital, while 40 participants were in the same status at a university hospital. One participant did not specify the type of institution.

## Data collection tools

### *The online exam*

Appointed faculty members from academic institutions across Türkiye, each with different areas of expertise in neurology, submitted their items to an online item bank organized into predefined subject area categories. The TBN members conduct meetings to review questions from the item bank that were appropriate for the subject area and made necessary revisions. From this item pool, TBN members selected the items for the 2023 exam through 20 h of online meetings and discussions to ensure content validity.

The exam consisted of two main sections. The first section included 70 MCQs with four to eight options with one correct answer. In this section, candidates were allowed to return to and modify their answers to questions they had already answered or left blank while answering the questions. The second section contained 10 KFPs, each with two to four items (totaling 30 items). Of the key feature items, 29 required short answers, while only one was a multiple-choice item that allowed for multiple options to be selected. In the KFP section, since additional information was provided in each question that could serve as a clue for answering the previous question, candidates were not allowed to return to or modify their answers to questions they had left blank or already answered. There were 100 items in total, each worth one point, with the maximum achievable score being 100 points.

Two standard-setting methods were applied to determine the exam's cut score: the Nedelsky method for MCQs and the Angoff method for KFPs.<sup>[13,14]</sup> Consequently, the cut score for the entire exam was estimated to be 65.0.

### *Feedback questionnaire for participants*

The questionnaire was developed by the researchers through a review of the relevant literature. Its content validity was evaluated by three experts in neurology, medical education, and measurement and evaluation. Modifications were made based on their feedback. The questionnaire comprised two sections. The first section included 14 structured items assessing participants' perspectives on the online exam. Participants rated these items on a 9-point Likert scale, ranging from 1 (strongly disagree/very poor) to 9 (strongly agree/very good). The second section

included nine items covering demographics, as well as one open-ended item for free-text feedback and additional remarks. Participants were asked to complete questionnaire through Microsoft Forms (Microsoft Corp., Redmond, WA, USA) immediately after the exam.

### **The procedure**

Explanations regarding the exam (location, date, time, item type, number of items, and information about the online exam platform to be used) were shared on the website of the Turkish Neurological Society two weeks before the exam. Detailed information was provided, specifically about Moodle, the online exam platform to be used for the exam. A practice exam was created on Moodle to help candidates become familiar with the exam interface. Like the actual exam, the practice exam included both MCQ and KFP items. Username and password information required to access the exam interface was sent to candidates' email addresses. Reminder emails were also sent to ensure that all candidates took the practice exam before the board exam.

The exam was conducted in December 2023 at one venue, under supervision, with online security measures such as individual passwords for login, question shuffling, and browser lockdown. The total exam time was 130 min, divided into MCQs (70 min), followed by KFPs (60 min). Each question was worth 1 point, with a maximum of 100 points obtainable. The MCQ scores were calculated automatically by the learning management system, while KFP responses were downloaded and scored by two different board members, with a final consensus mark determined through discussion among all board members. The final score was the sum of MCQ and KFP scores. Candidates' feedback was obtained via an online survey using a 9-point scale (1=strongly disagree/very bad; 9=strongly agree/very good).

### **Statistical analysis**

Data were analyzed using IBM SPSS version 25.0 software (IBM Corp., Armonk, NY, USA) was used for data analysis. The normal distribution of continuous variables was examined using the Shapiro-Wilk test ( $n < 50$ ) and the Kolmogorov-Smirnov test ( $n \geq 50$ ) and presented as mean  $\pm$  standard deviation (SD) and median (min-max). Spearman correlation analysis was used to investigate the relationship between numeric variables. Due to the nonnormal distribution of numeric variables, the comparison between the two groups was performed using

the Mann-Whitney U test. Categorical variables were presented as numbers and percentages. The relationship between categorical variables was examined using Pearson's chi-square and Fisher's exact test. A significance level of 0.05 was accepted for all hypotheses.

## RESULTS

The findings of the study are presented below in two stages: findings related to the exam and feedback of the participants.

### Findings related to the exam

Out of 96 applicants, a total of 60 participants (15 neurologists and 45 neurology residents in their final year of training who met the eligibility criteria) took the exam. The demographic characteristics of the exam participants are presented in Table 1. Following the review of appeals submitted by candidates after the exam, seven MCQs were omitted from the exam set due to various reasons. The omitted items were considered correctly answered by all candidates. Descriptive statistics summarizing the examination results are presented in Table 2.

As observed in Table 2, the mean scores were  $47.7 \pm 6.6$  out of 70 for the MCQs,  $16.1 \pm 3.8$  out of 30 for the KFPs, and  $63.8 \pm 8.8$  overall. Thirty (50.0%) candidates scored higher than the minimum acceptable level of performance (65/100) and passed the exam. The mean score percentage for the MCQ section (68.2%) was significantly higher than for the KFP section (53.1%;  $p < 0.001$ ).

The minimum score of the MCQ section was 32 out of 70, while it was 1.85 out of 30 for the KFP section. In other words, the minimum score percentages were estimated as 46% and 6% for

the MCQ and KFP sections, respectively. When examining the highest scores obtained in the exam (maximum score percentage), this value was 62 (89%) for the MCQ section and 22.9 (76%) for the KFP section. When examining the distribution of the scores, it was observed that the skewness and kurtosis values fell within normal limits for the MCQ section and the overall exam, while for the KFP section, the values deviated from normal distribution. This finding was supported by the normality test ( $p < 0.05$  for the KFP section).

Cronbach's alpha coefficients for the MCQ (63 items) and KFP (30 items) sections were estimated as 0.75 and 0.67, respectively.

### Feedback of the participants

A large proportion of participants ( $n=56$ , 93%) responded to the feedback questionnaire provided after the exam and shared their opinions. The 14 items included in the questionnaire and their mean scores out of nine are presented in Table 3.

According to the findings, the participants found both the MCQs and the KFPs to be quite difficult. Furthermore, candidates found the KFP section ( $6.6 \pm 1.8$ ) more challenging than the MCQ section ( $6.1 \pm 1.6$ ). The proportion of those who indicated that the clinical case questions were consistent with current clinical practice and believed that these questions assessed clinical problem-solving skills was above average. The highest level of satisfaction was reported for the inclusion of clinical case questions. It was observed that the items related to the adequacy of the exam duration and the organization of the exam received

**TABLE 2**  
Descriptive statistics

	MCQs	KFPs	Total exam
Number of participants	60	60	60
Mean	47.7	16.1	63.8
Median	48.5	16.8	64.5
Standard deviation	6.61	3.82	8.79
Mean score percentage	68.2	53.1	63.8
Minimum	32	1.85	42.5
Maximum	62	22.9	82.8
Skewness	0.509	-0.928	-0.492
Kurtosis	-0.00443	2.15	0.288
Shapiro-Wilk W	0.969	0.951	0.969
Shapiro-Wilk p	0.130	0.018	0.132

MCQs: Multiple-choice questions; KFPs: Key feature problems.

**TABLE 1**

Demographic characteristics of study participants ( $n=60$ )

Characteristic	n	%	Mean $\pm$ SD
Age (year)			30.7 $\pm$ 1.5
Sex			
Male	25	41.7	
Female	35	58.3	
Residency status			
Graduated	15	25.0	
In-training	45	75.0	
Experience in specialty			1.07 $\pm$ 1.44
Residency training			4.96 $\pm$ 0.54

SD: Standard deviation.

**TABLE 3**  
Distribution of the answers in the feedback questionnaire for candidates

Items	Mean
1. The multiple-choice questions were difficult.	6.13
2. The clinical case questions were difficult.	6.59
3. The clinical case questions were consistent with my current clinical practice.	6.00
4. I believe the clinical case questions measured my clinical problem-solving skills.	5.55
5. I appreciated the inclusion of clinical case questions in the exam.	6.63
6. The exam duration was sufficient.	4.46
7. The exam was well-organized.	4.70
8. The exam was designed to effectively differentiate between knowledgeable and less knowledgeable candidates.	5.32
9. The distribution of questions across topics was balanced throughout the exam.	5.64
10. The exam was suitable for assessment of neurologists/neurology residents.	5.39
11. The exam content was consistent with the scope of neurology residency training.	5.68
12. The exam venue was suitable in terms of lighting, temperature, sound, and other physical conditions.	6.02
13. It was easy to take the exam on a computer.	5.45
14. The computer interface used for the exam was user-friendly.	5.96

average scores, with the two lowest scores from the questionnaire being related to these items. The participants' opinions regarding the exam's discriminative ability, the balanced distribution of questions across topics, its suitability as an assessment tool for specialists, and its alignment with neurology residency training were also above average. Finally, candidates provided positive feedback, indicating that the exam venue was comfortable, the digital format was easy to use, and the exam user interface was user-friendly.

An open-ended question was also included in the questionnaire to allow participants to share their opinions on the overall exam process. Nineteen (32%) participants responded to this question. The feedback shared by participants regarding the overall exam can be summarized as follows.

Participants expressed satisfaction with the exam being conducted in a computer-based format (n=2) but preferred taking the exam in their own clinics under supervision, rather than traveling to another city for the exam. One participant, however, found the computer-based exam inefficient, cumbersome, and unnecessary. The inability to return to and modify answers or fill in blank responses for the KFP section was criticized by some candidates (n=6). One participant stated that this practice was not consistent with current neurology practice, while another mentioned that it unnecessarily made

the exam more difficult. Participants shared the difficulties they faced due to internet connectivity issues during the exam. They noted that trying to reconnect after a disconnection increased their anxiety (n=4). Candidates who felt that the overall exam time was insufficient (n=4) specifically mentioned that more time should be allocated to the KFP section. Due to the exam being conducted in a computer-based format for the first time, situations arose that required explanations during the exam. Participants mentioned that the noise generated by these explanations distracted them and caused difficulties in concentrating (n=8).

## DISCUSSION

This study highlights the successful digital transformation of the TNNBE, including its implementation and candidate feedback. The introduction of a computer-based exam facilitated streamlined administration and automatic scoring for MCQs, while still requiring manual scoring for KFPs. Exam performance data revealed higher scores for MCQs compared to KFPs, indicating that KFPs designed to assess clinical decisions were perceived as more challenging by candidates. On the other hand, candidate feedback was predominantly positive, praising the inclusion of clinical KFPs questions and the user-friendly interface, although some noted challenges with time constraints and technical issues during the exam.

The implementation of computer-based assessment (CBA) in medical education offers several advantages over traditional paper-based methods, including enhanced efficiency and reduced time required for administration and scoring.<sup>[15,16]</sup> Computer-based assessments provide higher quality and more detailed feedback compared to paper-based methods, and can minimize human errors in scoring.<sup>[16,17]</sup> Students' perceptions of CBA also improve over time, suggesting increased acceptance and comfort with the format and may reduce stress for examinees by allowing them to take exams in a familiar environment.<sup>[8,15,18]</sup>

The transition to CBA in our case demonstrated significant advantages, particularly in reducing the time required for administering and scoring MCQs. By integrating multiple question formats, including KFPs, CBA enabled a more comprehensive and nuanced evaluation of candidate competencies while maintaining operational efficiency. However, challenges such as technical issues and the manual scoring of KFPs highlight areas for further refinement. These findings, consistent with the literature, underscore the potential of CBA to streamline assessment processes while reflecting a consistent challenge in integrating complex clinical reasoning assessments into digital platforms.

Multiple-choice questions continue to serve as an optimal assessment tool for both low and high-stakes examinations.<sup>[3]</sup> The literature demonstrates that reliability scores increase with examination duration, ranging from 0.62 for a 1-h exam to 0.76 for a 2-h exam, and up to 0.93 for a 4-h exam.<sup>[19]</sup> In the literature, reliability scores exceeding 0.70 are generally considered indicative of a reliable measure, although some sources suggest aiming for a threshold above 0.80.<sup>[3]</sup> In this study, a reliability coefficient of 0.75 was obtained, consistent with the trends reported in the literature.

Previous studies reported KFP reliability coefficients ranging from 0.43 for a 1-h examination, 0.49 for a 2-h examination, to 0.67 for a 4-h examination.<sup>[20,21]</sup> A review study reported internal consistency reliability values ranging from 0.49 to 0.95, with the highest values observed when 25 to 40 cases were included.<sup>[22]</sup> In this study, the KFP reliability coefficient was 0.67, equivalent to the level typically achieved in a 4-h examination, despite the exam duration being only 1-h.

The limited available literature aligns with our study, reporting that participants found KFPs in examinations useful, engaging, and reflective of clinical practice, with a high level of acceptance and appreciation for the inclusion of this type clinical case-based questions.<sup>[20,21]</sup>

Online exams present multiple challenges for students, including heightened anxiety due to technological issues such as internet connectivity and unfamiliarity with the computer-based format, technological and infrastructural inadequacies disrupting the exam process, and insufficient time allocation for complex sections like KFPs, all of which negatively impact performance and concentration.<sup>[17,23-25]</sup> Consistent with these findings, participants in our study reported increased anxiety due to connectivity problems, insufficient exam time, particularly for the KFP section, and difficulties concentrating caused by distractions during the first-time implementation of the computer-based format.

Limitations of this study include a modest sample size and its single-center design, both of which constrain generalizability. Scoring of the KFPs required manual review because acceptable answers varied, limiting opportunities for automation. Technical issues—such as intermittent internet connectivity and environmental distractions—may have confounded the results. Participants also reported insufficient time, especially for the KFP section. Finally, the absence of comparative or longitudinal data prevents assessment of the digital format's long-term impact.

In conclusion, these findings underscore the benefits and challenges of transitioning to a digital exam format. They highlight the importance of addressing various logistical and scoring complexities that can arise in future iterations of computer-based examinations. This includes ensuring reliable internet connectivity, providing sufficient exam time, and minimizing distractions during the testing process to enhance the overall candidate experience and performance. Future research should focus on exploring the impact of digital exams on candidates' cognitive load, performance, and satisfaction across diverse educational settings. Additionally, studies investigating the long-term implications of computer-based testing on assessment validity, equity, and accessibility could provide valuable insights for optimizing digital examination frameworks.



**Acknowledgements:** The authors would like to thank the Turkish Neurological Society for their invaluable support in conducting this study. We are also grateful to the TNNBE participants for their time, commitment, and willingness.

**Data Sharing Statement:** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Author Contributions:** Were involved in the conceptualization of the research, development of the data collection tool, data collection, and interpretation of the results: S.A.C., G.T.T., H.M., Ö.E., F.G., F.E., S.G.K., F.F.E., M.D., B.N., E.S.; Contributed to the research design and took the lead in writing the manuscript: S.A.C., G.T.T. All authors provided critical feedback, helped shape the research, analysis, and manuscript, and reviewed and approved the final manuscript.

**Conflict of Interest:** The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

**Funding:** The authors received no financial support for the research and/or authorship of this article.

## REFERENCES

1. Çalışkan SA, Kansu T, Öztürk Ş, Araç N, Balkan S, Bora İ, et al. Turkish Neurology Board Examinations: 2004-2008. *Turk J Neurol* 2008;14:422-429.
2. Öztürk Ş. Turkish Neurology Board Examination - Process, Results and Evaluation. *Turk J Neurol* 2006;12:177-184.
3. Shumway JM, Harden RM; Association for Medical Education in Europe. AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Med Teach* 2003;25:569-84. doi: 10.1080/0142159032000151907.
4. Yudkowsky R, Soo Y, Downing SM. Assessment in health professions education editado por Rachel Yudkowsky, Yoon Soo Park, Steven M. Downing. Routledge, 2020. Available at: <https://www.routledge.com/Assessment-in-Health-Professions-Education/Yudkowsky-Park-Downing/p/book/9781315166902> [Accessed: Dec. 20, 2024]
5. Renes J, van der Vleuten CPM, Collares CF. Utility of a multimodal computer-based assessment format for assessment with a higher degree of reliability and validity. *Med Teach* 2023;45:433-41. doi: 10.1080/0142159X.2022.2137011.
6. van Wijk EV, Janse RJ, Ruijter BN, Rohling JHT, van der Kraan J, Crobach S, et al. Use of very short answer questions compared to multiple choice questions in undergraduate medical students: An external validation study. *PLoS One* 2023;18:e0288558. doi: 10.1371/journal.pone.0288558.
7. Schuwirth LW, van der Vleuten CP. Different written assessment methods: What can be said about their strengths and weaknesses? *Med Educ* 2004;38:974-9. doi: 10.1111/j.1365-2929.2004.01916.x.
8. Caliskan SA, Öztürk S. Implementing 'key-features problems' to Turkish neurology board examination. *J Neurol Sci* 2017;381:283. doi: 10.1016/j.jns.2017.08.807.
9. Medical Council of Canada. Guidelines for the Development of Key Feature Problems & Test Cases. Aug. 2012. Available at: <https://mcc.ca/wp-content/uploads/CDM-Guidelines.pdf> [Accessed: Dec. 21, 2024]
10. Bordage G, Gordon P. An Alternative to PMPs: The Key Features. Concept. Further Developments in Assessing Clinical Competence. 2nd Ottawa Conference; 1987. p. 59-75.
11. Page G, Bordage G. The Medical Council of Canada's key features project: A more valid written examination of clinical decision-making skills. *Acad Med* 1995;70:104-10. doi: 10.1097/00001888-199502000-00012.
12. Page G, Bordage G, Allen T. Developing key-feature problems and examinations to assess clinical decision-making skills. *Acad Med* 1995;70:194-201. doi: 10.1097/00001888-199503000-00009.
13. Cusimano MD. Standard setting in medical education. *Acad Med* 1996;71:S112-20. doi: 10.1097/00001888-199610000-00062.
14. Downing SM, Tekian A, Yudkowsky R. Procedures for establishing defensible absolute passing scores on performance examinations in health professions education. *Teach Learn Med* 2006;18:50-7. doi: 10.1207/s15328015tlm1801\_11.
15. Guimarães B, Ribeiro J, Cruz B, Ferreira A, Alves H, Cruz-Correia R, et al. Performance equivalency between computer-based and traditional pen-and-paper assessment: A case study in clinical anatomy. *Anat Sci Educ* 2018;11:124-36. doi: 10.1002/ase.1720.
16. Phillips AC, Mackintosh SF, Gibbs C, Ng L, Fryer CE. A comparison of electronic and paper-based clinical skills assessment: Systematic review. *Med Teach* 2019;41:1151-9. doi: 10.1080/0142159X.2019.1623387.
17. Heidarzadeh A. Opportunities and Challenges of Online Take-Home Exams in Medical Education. *J Med Edu* 2021;20: e112512. doi: 10.5812/JME.112512.
18. Grumer M, Brüstle P, Lambeck J, Biller S, Brich J. Validation and perception of a key feature problem examination in neurology. *PLoS One* 2019;14:e0224131. doi: 10.1371/journal.pone.0224131.
19. Norcini JJ, Swanson DB, Grosso IJ, Webster GD. Reliability, validity and efficiency of multiple choice question and patient management problem item formats in assessment of clinical competence. *Med Educ* 1985;19:238-47. doi: 10.1111/j.1365-2923.1985.tb01314.x.
20. Hatala R, Norman GR. Adapting the key features examination for a clinical clerkship. *Med Educ* 2002;36:160-5. doi: 10.1046/j.1365-2923.2002.01067.x.
21. Yılmaz Y, Çalışkan SA, Darcan Ş, Darendeliler F. Flipped learning in faculty development programs: opportunities for greater faculty engagement, self-learning, collaboration and discussion. *TJB* 2022;47:127-135. doi: 10.1515/TJB-2021-0071.

22. Hrynychak P, Takahashi SG, Nayer M. Key-feature questions for assessment of clinical reasoning: A literature review. *Med Educ* 2014;48:870-83. doi: 10.1111/medu.12509.
23. Bilge UN, Mehtap A, Cenk A. Measurement and evaluation challenges in distance education: Systematic review and qualitative meta-synthesis. *Journal of Educational Technology* 2023;19:1. doi: 10.26634/JET.19.4.19368.
24. Kolagari S, Modanloo M, Rahmati R, Sabzi Z, Ataee AJ. The effect of computer-based tests on nursing students' test anxiety: A Quasi-experimental study. *Acta Inform Med* 2018;26:115-8. doi: 10.5455/aim.2018.26.115-118.
25. Butler-Henderson K, Crawford J. A systematic review of online examinations: A pedagogical innovation for scalable authentication and integrity. *Comput Educ* 2020;159:104024. doi: 10.1016/j.compedu.2020.104024.