

Usability and preference of electronic vs. paper and pencil OSCE checklists by examiners and influence of checklist type on missed ratings in the Swiss Federal Licensing Exam

Abstract

Background: Only a few studies with small sample sizes have compared electronic Objective Structured Clinical Examination (OSCE) rating checklists with traditional paper-based OSCE rating checklists. In this study, the examiner-perceived usability and preference for type of OSCE checklist (electronic vs. paper based) were compared, and the influence of OSCE checklist type on missed ratings was determined, for the Swiss Federal Licensing Examination in clinical skills for human medicine.

Methods: All examiners in the Swiss Federal Licensing Examination in clinical skills for human medicine were invited over two subsequent years to evaluate the OSCE checklist type they had worked with during the examination. This was based on a questionnaire with 14 closed questions (i.e., demographic, checklist-type experience, perceived usability, checklist type preference). Furthermore, the numbers of missed ratings for the paper-based checklist were recorded.

Results: The data from these examiners ($n=377$) with experience of both OSCE checklist types were analyzed. The electronic OSCE checklist was rated significantly higher on all usability aspects (i.e., ease of use, candidate rating and error correction, clarity, distraction using the checklist, overall satisfaction), except for the speed of registering comments (no significant difference). The majority of the examiners in both years (2014: 54.5%, $n=60$, 2015: 89.8%, $n=230$) reported preference for working with the electronic OSCE checklist in the future. Missed ratings were seen for 14.2% of the paper-based OSCE checklists, which were prevented with the electronic OSCE checklists.

Conclusions: Electronic OSCE checklists were rated significantly more user-friendly and were preferred over paper-based OSCE checklists by a broad national sample of examiners, supporting previous results from faculty-level examinations. Furthermore, missed ratings were prevented with the electronic OSCE checklists. Overall, the use of electronic OSCE checklists is therefore advisable.

Keywords: OSCE, checklists, electronic, usability, evaluation, national

Felicitas L. Wagner¹

Sabine Feller¹

Felix M. Schmitz²

Philippe G.

Zimmermann²

Rabea Krings¹

Sissel Guttormsen³

Sören Huwendiek¹

1 University of Bern, Institute for Medical Education, Department for Assessment and Evaluation, Bern, Switzerland

2 University of Bern, Institute for Medical Education, Department for Software Development, Usability Consulting and IT Infrastructure, Bern, Switzerland

3 University of Bern, Institute for Medical Education, Bern, Switzerland

Background

The administration of Objective Structured Clinical Examinations (OSCEs) [1] has become the standard method to reliably assess practical skills in medical education [2], [3], [4], [5]. In an OSCE, the clinical and communication skills of the students are probed while they complete multiple stations, each of which requires them to work on a particular medical problem. During the OSCE, their performance is observed and rated by examiners through standardized, structured checklists.

Traditionally, paper-based checklists have been used to rate the candidates during an OSCE, while more recently, with the emergence of mobile electronic devices, electron-

ic OSCE checklists have been developed and successfully implemented [6], [7], [8]. Electronic OSCE checklists are usually presented on “tablet” computers, as these are more flexible than laptops or desktop devices [7]. To date, however, few studies have compared electronic vs. paper-based OSCE checklists regarding their usability and the examiner checklist type preference. As examiners are an important stakeholder group and the users of the checklists, it is important what their usability perceptions and preferences are. According to Petersen [9], “the underlying construct of usability is to make objects compatible” with the user. Objects can be hardware, software, or every-day tools. Their design should not hinder the completion of the tasks they were designed for. Petersen

further describes usability as how easy to use an object is and underlines the importance of the quality of the generated output for which an object has been created to determine its usability. OSCEs create a high cognitive load [10] for the examiners. Examiners have to simultaneously observe and rate the performances of the candidates, find the respective items in the OSCE checklist, decide on the amount of points the candidates receive for the different checklist items, and also handle the OSCE checklist. Paper-based OSCE checklists, e.g., can be several pages long [6]. Furthermore, the examination times are usually long. Considering these aspects, the OSCE checklists need to be easy to use and allow the examiners to rate candidates in an efficient and effective way to not increase the cognitive load and hinder the completion of the task they were created for, that is, rating the performances of the candidates. It is therefore important that they do not distract the examiners from their observing and rating of the performances of the candidates during examinations. Finally, also the output quality is an important aspect in the context of usability [9]. Paper-based OSCE checklists can have specific issues regarding the quality of the data. These often contain a substantial number of missed ratings, as checklist items are easily overlooked during an examination, which makes time-consuming manual data verification necessary [6]. Furthermore, paper-based checklists usually have to be scanned to transfer the OSCE results into a digital form, as a necessary step to make the results available for analysis. The use of electronic OSCE checklists on the other hand leads to better data quality because missed ratings can be prevented. Furthermore, the data can be downloaded and stored electronically directly following the examination, making the preparation of the data for analysis less error-prone and time-consuming compared to paper-based OSCE checklists [6].

A systematic review in 2019 [11] mentioned three studies that investigated examiner perceptions of electronic OSCE checklists, although only two of these [6], [12] investigated the perceived usability of paper-based versus electronic OSCE checklists. Along with studies by Hochlehnert et al. [7] and Currie et al. [8], these two studies [6], [12] suggested that the perceived usability of the respective electronic OSCE checklist tools that were used was high. However, these involved only small to medium numbers of examiners ($n=10$ [6], $n=35$ and 33 [7], $n=93$ [8], $n=43$ [12]) who were from a single faculty, thus limiting the generalizability of their results. Furthermore, as described above, one critical issue with paper-based OSCE checklists are missed ratings. However, to date little is known about the size of this problem. To our knowledge, only three studies [6], [8], [13] have addressed this problem to date.

Our research questions therefore were:

1. How does a large group of examiners from five different faculties, experienced with paper-based and electronic OSCE checklists, rate the usability of elec-

tronic versus paper-based OSCE checklists in a national licensing skills examination?

2. Which type of checklist (electronic vs. paper-based) is preferred by the examiners of a national licensing skills examination?
3. How large is the amount of missed ratings in the paper-based OSCE checklists?

Methods

Setting

The Swiss Federal Licensing Examination in clinical skills (which follows the principles of the OSCE examination) contained 12 different OSCE stations in both 2014 and 2015. The examination was conducted over three consecutive days at five faculties of Human Medicine in Basel, Bern, Geneva, Lausanne and Zurich.

The paper-based OSCE checklists that were used for the 2014 examination were created using a teleform information capture system (Cardiff Software). These checklists had to be scanned after the examination, to make the results available for analysis (DR-6050 C image formula scanner; Canon). The paper-based OSCE checklists contained between 25 and 38 items, which was due to the different contents of the individual stations. These items were always sorted according to the same four dimensions, in the following order:

1. history taking,
2. physical examination,
3. management, and
4. communication.

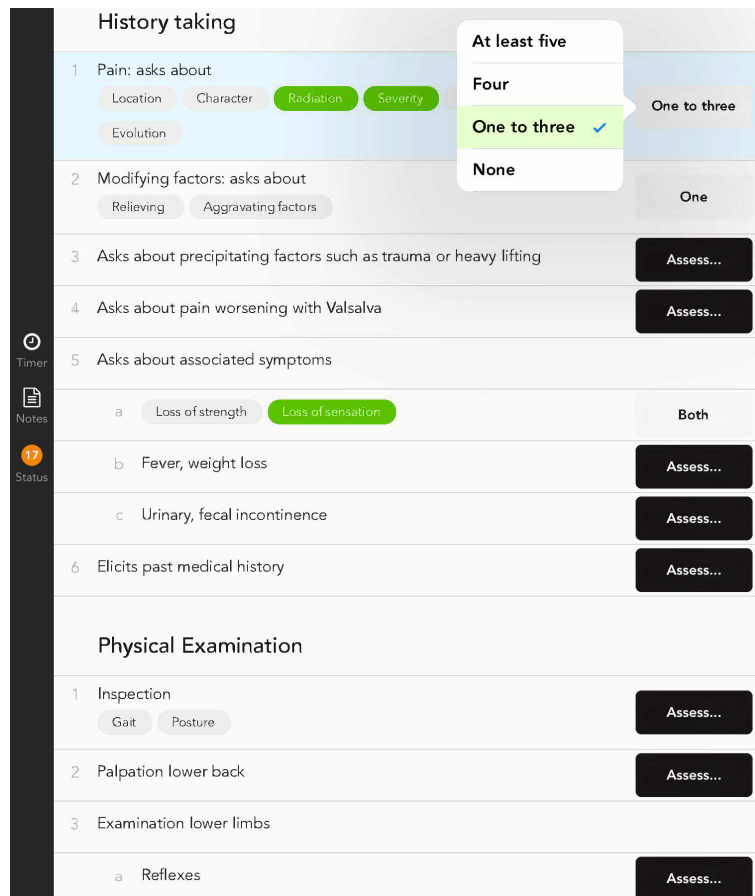
The paper-based OSCE checklists further contained two global ratings for overall impression. At the top of the paper-based OSCE checklists, a reminder sentence was included stating that all of the items in the checklist had to be completed.

The electronic OSCE checklists in the 2015 examination were completed using identical fourth generation iPad devices (2014, Apple). The iPad app "OSCE-Eval" [<https://eosce.ch/>], version 2.1, was used to rate the candidates. This app includes different features to reduce potential errors and to ensure that the collected data are readily available for subsequent analyses. The users receive visual feedback to make sure that every item is completed. Every item had to be evaluated to finally submit the checklist. After the final assessment, no more changes were possible. Encrypted data were synchronized with a secure server, and the data were automatically exported to an electronic spreadsheet. Depending on the specific content of the stations, the electronic OSCE checklists contained between 21 and 42 items. The items were sorted according to the same four dimensions as in the paper-based OSCE checklist. In addition, there were two global ratings in each checklist for overall impressions (see figure 1 and figure 2 for examples of both checklist types). The electronic checklists were developed accord-

fill in the appropriate fields as shown here:  clearly mark corrections:  please mark **every** criterion!

	Yes	+/-	No
History taking			
A1 Pain: asks about location, character, radiation, severity, duration, onset, evolution. 5-7 = yes, 4 = +/-	☐	☐	☐
A2 Modifying factors: asks about relieving factors, aggravating factors. 2 = yes, 1 = +/-	☐	☐	☐
A3 Asks about precipitating event such as trauma or heavy lifting. Yes / no	☐		☐
A4 Asks about pain worsening with Valsalva. Yes / no	☐		☐
A5 Asks about associated symptoms: a) Loss of strength, loss of sensation 2=yes, 1=+/- b) Fever, weight loss. 1 = yes c) Urinary, fecal incontinence. 1 = yes	☐ ☐ ☐	☐	☐ ☐ ☐
A6 Elicits past medical history. Yes / no	☐		☐
Physical examination			
S1 Inspection: gait, posture. 2 = yes, 1 = +/-	☐	☐	☐
S2 Palpation lower back. Completed satisfactorily / Attempted only / Not done	☐	☐	☐
S3 Examination lower limbs: a) Reflexes. Completed satisfactorily / Attempted only / Not done	☐		☐

Figure 1: Example image of part of the paper-based OSCE checklist (fictitious candidate)



History taking

- Pain: asks about
 - Location
 - Character
 - Radiation
 - Severity
 - Evolution
- Modifying factors: asks about
 - Relieving
 - Aggravating factors
- Asks about precipitating factors such as trauma or heavy lifting
- Asks about pain worsening with Valsalva
- Asks about associated symptoms
 - Loss of strength
 - Loss of sensation
 - Fever, weight loss
 - Urinary, fecal incontinence
- Elicits past medical history

Physical Examination

- Inspection
 - Gait
 - Posture
- Palpation lower back
- Examination lower limbs
 - Reflexes

Figure 2: Example image of part of the electronic OSCE checklist (fictitious candidate)

ing to the usability standards at the time, and considerable effort was invested to iteratively check their acceptance during their development.

In the introduction to the OSCE exam in both 2014 and 2015, the examiners were briefly instructed on how to use the checklists and how to rate the candidates. Additionally, the examiners working with the electronic OSCE checklist (in 2015) had the possibility to watch a specific, 6-min training video before the examination, to familiarize themselves with the checklist and the software. The

training video was the same that was used by Schmitz et al. [6], who reported that the video was comprehensible and effective.

Sample

In 2014, 696 examiners rated the performances of the candidates at the different stations using the paper-based OSCE checklist. In 2015, 696 examiners rated the performances of the candidates at the different stations

Table 1: Evaluation questionnaire. The questions have been translated from German and French to English for clarity here.

Usability		
1	The checklist was easy to use.	do not agree at all (1) – fully agree (7)*
2	The structure of the checklist was clear.	do not agree at all (1) – fully agree (7)*
	The checklist enables...	
3	Fast rating of candidates	do not agree at all (1) – fully agree (7)*
4	Fast registration of comments	do not agree at all (1) – fully agree (7)*
5	Fast correction of input errors	do not agree at all (1) – fully agree (7)*
6	Filling in the checklist distracted me from observing the candidates.	do not agree at all (1) – fully agree (7)*
7	Overall I am satisfied with the checklist.	do not agree at all (1) – fully agree (7)*
Demographics		
8	Faculty	Basel / Bern / Geneva / Lausanne / Zurich
9	Gender	male / female
10	Age	<35 years / 36-40 years / 41-50 years / 51-60 years / >60 years
11	How many times including today have you worked with paper-based OSCE checklists to rate candidates?	Never / once / twice / three times / four times / more than four times
12	How many times including today have you worked with electronic OSCE-checklists to rate candidates?	Never / once / twice / three times / four times / more than four times
13	How experienced are you with touch screen devices (tablets, smartphones, etc.)?	Not experienced at all (1) – very experienced (7)
Preference of checklist type:		
14	If you have already worked with both checklist types of checklist: what type would you prefer to work with in the future?	Paper-based checklist / electronic checklist / no preference

Note: *Only the extreme points of the scale were labeled.

using the electronic OSCE checklist. All examiners were invited to take part in the survey. Only data from examiners who had worked with both checklist types (paper-based and electronic) at least one time have been included to allow for comparisons between the electronic and the paper-based OSCE checklist.

Material

The examiners evaluated the OSCE checklist type (i.e., paper-based in 2014, electronic in 2015) that they had used to rate the candidates through completion of a questionnaire (see table 1). The questionnaire contained a total of 14 closed questions: demographic information (three questions); experience with touch-screen devices and the two different OSCE checklist types (electronic, paper-based; three questions); subjective usability (seven questions); and future checklist type preference (one question). The replies were based on a 7-point Likert scale, from 1, as “do not agree at all”, to 7, as “fully agree”.

The usability questions were informed by System Usability Scale [14]. However, the questionnaire had only seven questions regarding usability, to reduce the efforts of the examiners, and six further questions concerning the specific requirements in an OSCE examination setting (e.g., questions regarding speed of rating the candidates, distraction from observing the candidates, speed of cor-

recting input errors). These questions further reflect usability aspects as indicated by Petersen [9] (ease of use, not hindring the completion of the task).

The questionnaire was available in German and French. To ensure that the questions were clear, think-aloud trials [15] with three examiners were conducted in both languages. During each think-aloud, the participants were asked to verbalize their thoughts while they were working on the questionnaire. This method thus allowed the identification of ambiguities and misconceptions, and it was used to make sure that the questions were correctly understood by the participants. The same questionnaire was used in 2014 and 2015.

Finally, to determine the completeness of the data, the number of missed ratings for the paper-based checklists was recorded.

Procedure

For both years, within the first week following the examination, all of the examiners were invited to evaluate the OSCE checklist type they had worked with during the examination (i.e., paper-based in 2014, electronic in 2015). Hence, a comparison across the two OSCE checklist types was possible. In 2014, the examiners received the questionnaire in paper-based form directly after the exam, in 2015 the questionnaire was implemented online using the online survey tool “Unipark” [https://

www.unipark.com/] and the examiners were invited via e-mail. This type of study was regarded as exempt from formal ethical approval according to the regulations of the Ethics Committee ("Kantonale Ethikkommission" Canton of Bern) associated with the Medical Faculty of the University of Bern. Participation in the survey was voluntary and anonymous and participants have not been exposed to any conceivable risks by participating in this study.

Statistical analysis

Frequencies were analysed to report demographics of the sample. For comparisons between the paper-based and the electronic OSCE checklist, Mann-Whitney-Tests were used as the data are ordinally scaled. The alpha level was adjusted for multiple testing according to Bonferroni. The effect size is given by Pearson's r , the magnitude of the effects was judged according to Cohen [16].

Results

Participants

In 2014, 696 examiners rated the performances of the candidates at the different stations using the paper-based OSCE checklist. Of these, 540 (78%) completed the survey questionnaire. In 2015, 696 examiners rated the performances of the candidates at the different stations using the electronic OSCE checklist. Here, the response rate was 58%, where 406 of the 696 examiners completed the survey questionnaire.

The questionnaire datasets from 35 (4%) examiners were removed from the database due to missing data (2014: 19 datasets; 2015: 16 datasets). All datasets that were removed contained no useful information as these persons stopped filling out the questionnaire after the language selection right in the beginning of the questionnaire. Datasets from 911 (96%) examiners were taken forward for the analysis. Of these 911 examiners, 377 (41%) reported that they had worked with both checklist types (paper-based and electronic) at least once. This included examiners from both years, as examiners who worked with the paper-based OSCE checklist in 2014 might have had previous experience with electronic OSCE checklists for other examinations. Only the responses from these examiners who had experience with both checklist types were included in the quantitative analysis. Table 2 details the experiences of the examiners for both of the checklist types.

Demographic questions

The demographics of the examiners in the final sample are presented in table 2.

Touch-screen experience

Most examiners from both years had vast ($M_{2014}=6.21$, $M_{2015}=6.08$; scale: 1, "not experienced at all", to 7, "very experienced") and comparable experience with the use of touch-screens ($U=13933.5$, $n.s.$). Thus, the examiner ratings and checklist type preferences (i.e., paper-based vs. electronic) are unlikely to be influenced by their prior experience with touch-screen devices.

Usability

Mann-Whitney-Tests indicated that the majority of the examiner's usability ratings significantly differed between the electronic and the paper-based OSCE checklists (see table 3). The Bonferroni adjusted alpha level for the seven comparisons described in the following was .007. The electronic OSCE checklist was rated as significantly easier to use ($U=9799.5$, $p<0.007$, $r=.33$) and clearer ($U=11441.5$, $p<0.007$, $r=.21$) than the paper-based OSCE checklist. The electronic OSCE checklist also scored significantly higher for speed of rating of candidates ($U=11408$, $p<0.007$, $r=.20$) and correcting input errors ($U=7766.5$, $p<0.007$, $r=.41$). Furthermore, filling in the electronic OSCE checklist was rated as significantly less distracting than filling in the paper-based OSCE checklist ($U=11673$, $p<0.007$, $r=.18$). The speed of writing comments was not rated differently between the two checklist types ($U=12554$, $p=0.035$, $n.s.$, $r=.11$). The overall satisfaction of examiners was also significantly higher with the electronic than the paper-based OSCE checklist ($U=7499$, $p<0.007$, $r=.42$) (see table 4). The effects sizes according to Cohen [16] were either medium (between 0.3 and 0.5), or low (between 0.1 and 0.3). Taken together, the electronic OSCE checklist received higher usability ratings than the paper-based OSCE checklist.

Checklist type preference

At the end of the questionnaire, the examiners from both 2014 and 2015 who had indicated experience with both electronic and paper-based OSCE checklists were asked whether they would rather work with electronic or paper-based OSCE checklists in the future (see table 4). Across both years, the majority of these examiners who had had experience with both electronic and paper-based OSCE checklists preferred the electronic format.

Missed ratings for the paper-based OSCE checklist

The number of missed ratings in the paper-based checklist was recorded (see table 5). In 1428 (14.2%) of all of the paper-based checklists, there were missed ratings. These checklists contained between 1 and 12 missed ratings ($M=1.45$, $SD=1.03$). In the majority of these checklists, one rating was missed.

Table 2: Demographics of the final study sample (examiners experienced with both the electronic and the paper-based OSCE checklist types)

Demographic	Detail	Examiners [n (%)]		
		2014	2015	Total
Faculty	Basel	32 (28.1)	60 (22.8)	92 (24.4)
	Bern	1 (0.9)	54 (20.5)	55 (14.6)
	Geneva	30 (26.3)	49 (18.6)	79 (21.0)
	Lausanne	27 (23.7)	30 (11.4)	57 (15.1)
	Zurich	24 (21.1)	70 (26.6)	94 (24.9)
Gender	Male	79 (69.3)	182 (69.2)	261 (69.2)
	Female	33 (28.9)	81 (30.8)	114 (30.2)
	Missing	2 (1.8)	0 (0.0)	2 (0.5)
Age group	<35 years	15 (13.2)	4 (1.5)	19 (5.0)
	35-40 years	17 (14.9)	50 (19.0)	67 (17.8)
	41-50 years	41 (36.0)	99 (37.6)	140 (37.1)
	51-60 years	29 (25.4)	85 (32.3)	114 (30.2)
	>60 years	12 (10.5)	24 (9.1)	36 (9.5)
	Missing	0 (0.0)	1 (0.4)	1 (0.3)
OSCE checklist type experience				
Paper based	Once	17 (14.9)	48 (18.3)	65 (17.2)
	Twice	21 (18.4)	53 (20.2)	74 (19.6)
	3 times	25 (21.9)	50 (19.0)	75 (19.9)
	4 times	12 (10.5)	33 (12.5)	45 (11.9)
	>4 times	39 (34.2)	79 (30.0)	118 (31.3)
Electronic	Once	73 (64.0)	145 (55.1)	218 (57.8)
	Twice	22 (19.3)	64 (24.3)	86 (22.8)
	3 times	9 (7.9)	23 (8.7)	32 (8.5)
	4 times	5 (4.4)	11 (4.2)	16 (4.2)
	>4 times	5 (4.4)	20 (7.6)	25 (6.6)

Table 3: Usability ratings of the electronic and paper-based OSCE checklists

Usability factor	Paper-based checklist						Electronic checklist						p
	n	M	SD	Mdn	Min	Max	n	M	SD	Mdn	Min	Max	
Ease of use	114	6.13	1.02	6	2	7	262	6.69	0.61	7	4	7	*
Clarity of checklist	113	6.16	0.96	6	2	7	263	6.50	0.83	7	3	7	*
Ease of rating candidates	113	5.97	1.00	6	2	7	263	6.32	0.94	7	2	7	*
Speed of writing comments	112	5.62	1.32	6	1	7	259	5.19	1.63	5	1	7	n.s.
Ease of correcting errors	111	5.36	1.62	6	1	7	260	6.55	0.86	7	2	7	*
Distraction from candidate assessment ¹	114	3.06	1.85	2	1	7	261	2.48	1.81	2	1	7	*
Overall satisfaction	113	5.18	1.60	6	1	7	260	6.39	0.93	7	1	7	*

Note: n=number of responses, M=mean, SD=standard deviation, Mdn=median, Min=minimal score, Max=maximal score; Scale from 1="do not agree at all" to 7="fully agree"; n.s. = not significant, *p<0.007

¹ lower numbers indicate a better score.

Table 4: Future preference for use of the OSCE checklist type

Checklist type	Examiners [n (%)]	
	2014	2015
Paper-based checklist	20 (18.2)	3 (1.2)
Electronic checklist	60 (54.5)	230 (89.8)
No preference	30 (27.3)	23 (9.0)

Table 5: Number of missings entries for the paper-based OSCE checklists

Number of missing entries	Checklists	
	<i>n</i>	%
0	8612	85.8
1	1067	10.6
2	230	2.3
3	59	0.6
4	30	0.3
5	29	0.3
6	4	0.04
7	4	0.04
8	2	0.02
9	1	0.01
10	1	0.01
12	1	0.01

Discussion

In this study that included a large sample of examiners experienced with electronic and paper-based OSCE checklists from five different faculties in the Swiss Federal licensing examination over two subsequent years, the perceived usability was significantly higher for the electronic than the paper-based OSCE checklist in all aspects except for the speed of writing comments, where no significant difference was observed. Furthermore, the vast majority of the examiners preferred to work with the electronic OSCE checklist in the future. Finally, there were missed ratings in 14.2% of all of the paper-based checklists versus none in the electronic checklist (as the electronic system required all checklist items to be completed).

We were able to show here that electronic OSCE checklists have higher perceived usability than paper-based checklists across a large dataset of examiners from five different faculties in the Swiss Federal Licensing examination. To date, the usability of electronic OSCE checklists has only been investigated in studies that have included mostly small to medium numbers of examiners [6], [7], [8], [12]. Given the high cognitive load [10] that these examiners experience during an OSCE, the usability of the OSCE checklist type to rate the candidates is essential. The examiners in this study overall were more satisfied with the electronic than with the paper-based OSCE checklist. They perceived the electronic OSCE checklist as clearer and easier to use compared to the paper-based OSCE checklist, while they also felt less distracted during the examinations. Furthermore, rating the candidates as well as correcting input errors was reported to be easier for the electronic compared to the paper-based OSCE checklist. However, there was no significant difference regarding the perceived speed of writing comments on a touch-pad (i.e., the electronic OSCE checklist) and writing comments on paper. It can be assumed that all of these factors contribute to allow the examiners to better focus their attention on the candidates and their performances

during the examinations, thus keeping the cognitive load of the examiners lower in the electronic OSCE checklist. Our analysis of missed ratings in the paper-based OSCE checklists showed that in some of the individual checklists, there were indeed a large number of missed items. However, overall, the amount of missed items was moderate, as this was seen for 14.2% of all of the paper-based OSCE checklists. In the majority of these (10.6% of all checklists containing missed items), only one missed item was found. The mean number of missed ratings in these checklists was 1.45 overall, which is a good deal lower compared to the missing ratings reported by Schmitz et al. [6]. The number of missed ratings for such paper-based checklists therefore appears to vary greatly between different examinations and examiners. With the use of electronic OSCE checklists on the other hand, missed ratings can be prevented altogether, which represents an important advantage of the electronic OSCE checklist, as it ensures higher data quality [6], [8], [12]. The output quality as an important aspect of usability [9] thus is clearly higher in electronic OSCE checklists. Despite the introduction to the paper-based OSCE checklist at the beginning of the exam and a reminder sentence on top of the checklist, there were still cases where examiners missed to rate items. One possible explanation could be that because of the lower usability of the paper-based checklist, the cognitive load became too high for the examiners in some situations which may have lead to the missed ratings.

The analysis of the checklist type preferences showed a clear result. Both in 2014 (54.5%) and in 2015 (89.9%) the majority of the examiners who were experienced with both OSCE checklist types specified that they would prefer to work with electronic OSCE checklists in the future, thus confirming the results of Schmitz et al. [6] and Currie et al. [8]. We assume that this finding is related to the higher perceived usability of the electronic OSCE checklist. The strengths of the present study are the large numbers of examiners involved, the high proportion of these examiners who had experience with both types of OSCE checklists (i.e., paper-based and electronic), and the

multi-institutional (national) setting that included examiners from five different faculties from two language regions.

A limitation of this study is that except for the missed ratings, these data are based on the subjective impressions of the examiners with regard to the use of these checklists, rather than providing an objective measure of usability.

Conclusions

The results of the present study show that the perceived usability of electronic OSCE checklists compared to traditional paper-based checklists was significantly higher, except for the speed of writing comments, where there was no statistically significant difference. The majority of the examiners experienced with both electronic and paper-based OSCE checklists preferred to work with the electronic OSCE checklists in the future. Importantly, while missed ratings are a common issue in paper-based checklists, they can be completely prevented when electronic OSCE checklists are used, which ensures high data quality. Overall, the results of the present study with a large number of examiners across different faculties show that if developed with a focus on usability, electronic OSCE checklists have advantages that go beyond technical factors, and so they can be considered as an advisable alternative to paper-based OSCE checklists.

Declarations

Ethics approval and consent to participate

This type of study was regarded as exempt from formal ethical approval according to the regulations of the Ethics Committee ('Kantonale Ethikkommission' Canton of Bern) associated with the Medical Faculty of the University of Bern. Nevertheless, we confirm that this study was conducted according to the Declaration of Helsinki where applicable for this type of study, and that participants took part on a voluntary basis, they cannot be identified by the material presented, and they have not been exposed to any conceivable risks by participating in this study.

Consent for publication

Not applicable. The images provided in this manuscript show a fictitious example created specifically for this publication to illustrate the electronic and paper-based OSCE checklists.

Availability of data and materials

The datasets analysed during the current study are available from the corresponding author on reasonable request. More information about the OSCE-Eval application can be found at the website [<https://eosce.ch/>].

Funding

This study was funded by the Institute for Medical Education, University of Bern, Switzerland.

Authors' contributions

F.W. analyzed and interpreted the data and wrote the manuscript with support from S.H. R.K. helped in the analysis of the data. S.H., S.F., F.S. and P.Z. constructed the questionnaire for the survey. S.G., S.F., R.K., F.S., P.Z. and S.H. critically revised the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

We would like to thank: the examiners for participating in the survey; the site responsables, Monica Escher, Felix Eymann, Claudia Glauser, David Gachoud and Roger Kropf; the members of the clinical skills working group, Gianmarco Balestra, Christoph Berendonk, Sylvie Félix, Sabine Feller, Philippe Huber, Matteo Monti, Kai Schnabel, and Ernst Jünger; and the IT support team of the Swiss Federal Licensing Examination at the faculties of Basel, Bern, Geneva, Lausanne and Zurich. Furthermore, we thank Daniel Stricker for his advice on the analysis of the missed items. Finally, we would like to thank the Federal Office of Public Health and the Examinations Commission for supporting this examination.

Competing interests

The authors declare that they have no competing interests.

The OSCE-Eval application that was used to implement the electronic OSCE checklist was designed and developed by the Department for Software Development, Usability Consulting and IT Infrastructure (ASCII), at the Institute for Medical Education. More information about the application can be accessed through the website [<https://eosce.ch/>].

References

1. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *Br Med J*. 1975;1(5955):447-451. DOI: 10.1136/bmj.1.5955.447
2. Barman A. Critiques on the objective structured clinical examination. *Ann Acad Med Singap*. 2005;34(8):478-482.
3. Reznick RK, Smee S, Baumber JS, Cohen R, Rothman A, Blackmore D, Berard M. Guidelines for estimating the real cost of an objective structured clinical examination. *Acad Med*. 1993;68(7):513-517. DOI: 10.1097/00001888-199307000-00001
4. Smee S. Skill based assessment. *Br Med J*. 2003;326(7391):703-706. DOI: 10.1136/bmj.326.7391.703

5. Zayyan M. Objective Structured Clinical Examination: The Assessment of Choice. *Oman Med J*. 2011;26(4):219-222. DOI: 10.5001/omj.2011.55
6. Schmitz FM, Zimmermann PG, Gaunt K, Stolze M, Guttormsen Schär S. Electronic Rating of Objective Structured Clinical Examinations: Mobile Digital Forms Beat Paper and Pencil Checklists in a Comparative Study. In: Holzinger A, Simoncic KM, editors. *Information Quality in eHealth*. Berlin: Springer; 2011. p.501-512. DOI: 10.1007/978-3-642-25364-5_35
7. Hochlehnert A, Schultz JH, Möltner A, Timbil S, Brass K, Jünger J. Electronic acquisition of OSCE performance using tablets. *GMS Z Med Ausbild*. 2015;32(4):Doc41. DOI: 10.3205/zma000983
8. Currie GP, Sinha S, Thomson F, Cleland J, Denison AR. Tablet computers in assessing performance in a high stakes exam: opinion matters. *J R Coll Physicians Edinb*. 2017;47(2):164-167. DOI: 10.4997/JRCPE.2017.215
9. Petersen D. Usability theory, practice and evaluation for learning objects. In: Koohang A, Harman K, editors. *Learning objects: Applications, implications, & future directions*. Informing Science; 2007. p.337-370.
10. Young JQ, Van Merriënboer J, During S, Ten Cate O. Cognitive load theory: implications for medical education: AMEE Guide No. 86. *Med Teach*. 2014;36(5):371-384. DOI: 10.3109/0142159X.2014.889290
11. Philips A, Fordyce Mackintosh S, Gibbs C, Ng L, Fryer CE. A comparison of electronic and paper-based clinical skills assessment: Systematic review. *Med Teach*. 2019;41(10):1151-1159. DOI: 10.1080/0142159X.2019.1623387
12. Treadwell I. The usability of personal digital assistants (PDAs) for assessment of practical performance. *Med Educ*. 2006;40(9):855-861. DOI: 10.1111/j.1365-2929.2006.02543.x
13. Denison A, Bate E, Thompson J. Tablet versus paper marking in assessment: feedback matters. *Perspect Med Educ*. 2016;5(2):108-113. DOI: 10.1007/s40037-016-0262-8
14. Brooke J. SUS-A quick and dirty usability scale. In: Jordan PJ, Thomas B, Weerdmeester BA, McClelland IL, editors. *Usability evaluation in industry*. London: Taylor & Francis; 1996. p.4-7.
15. Collins D. Pretesting survey instruments: An overview of cognitive methods. *Qual Life Res*. 2003;12(3):229-238. DOI: 10.1023/A:1023254226592
16. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Erlbaum; 1988.

Corresponding author:

Felicitas L. Wagner
University of Bern, Institute for Medical Education,
Department for Assessment and Evaluation, Mittelstr.
43, CH-3012 Bern, Switzerland
felicitas.wagner@iml.unibe.ch

Please cite as

Wagner FL, Feller S, Schmitz FM, Zimmermann PG, Krings R, Guttormsen S, Huwendiek S. Usability and preference of electronic vs. paper and pencil OSCE checklists by examiners and influence of checklist type on missed ratings in the Swiss Federal Licensing Exam. *GMS J Med Educ*. 2022;39(2):Doc24. DOI: 10.3205/zma001545, URN: urn:nbn:de:0183-zma0015453

This article is freely available from

<https://doi.org/10.3205/zma001545>

Received: 2021-01-13

Revised: 2022-01-28

Accepted: 2022-02-09

Published: 2022-04-14

Copyright

©2022 Wagner et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.

Usability und Präferenz von elektronischen OSCE-Checklisten im Vergleich zu papierbasierten Checklisten gemäss Prüfenden und Einfluss des Checklisten-Typs auf fehlende Bewertungen in der Eidgenössischen Prüfung

Zusammenfassung

Hintergrund: Nur wenige Studien mit kleinen Stichprobengrößen haben elektronische OSCE-Checklisten (Objective Structured Clinical Examination) mit traditionellen OSCE-Checklisten in Papierform verglichen. In dieser Studie wurden die von Prüfenden wahrgenommene Usability und Präferenz für den OSCE-Checklisten-Typ (elektronisch vs. papierbasiert) verglichen und der Einfluss des OSCE-Checklisten-Typs auf fehlende Bewertungen ermittelt für die Eidgenössische Clinical Skills-Prüfung Humanmedizin in der Schweiz.

Methode: Die Prüfenden der Eidgenössischen Clinical Skills-Prüfung Humanmedizin wurden in zwei aufeinanderfolgenden Jahren gebeten, den OSCE-Checklisten-Typ zu bewerten, mit dem sie während der Prüfung gearbeitet hatten. Dies geschah anhand eines Fragebogens mit 14 geschlossenen Fragen (demographische Angaben, Erfahrung mit dem Checklisten-Typ, wahrgenommene Usability, Präferenz für den Checklisten-Typ). Außerdem wurde die Anzahl der fehlenden Bewertungen bei der papierbasierten Checkliste erfasst.

Resultate: Die Daten derjenigen Prüfenden (n=377) mit Erfahrung mit beiden OSCE-Checklisten-Typen wurden ausgewertet. Die elektronische OSCE-Checkliste wurde bei allen Aspekten der Usability (einfache Benutzung, Kandidierendenbewertung und Fehlerkorrektur, Übersichtlichkeit, Ablenkung bei der Verwendung der Checkliste, Gesamtzufriedenheit) signifikant besser bewertet, mit Ausnahme der Geschwindigkeit des Erfassens von Kommentaren (kein signifikanter Unterschied). Die Mehrheit der Prüfenden in beiden Jahren (2014: 54.5%, n=60, 2015: 89.8%, n=230) gab an, in Zukunft lieber mit der elektronischen OSCE-Checkliste arbeiten zu wollen. Bei 14.2% der papierbasierten OSCE-Checklisten wurden fehlende Bewertungen festgestellt, welche mit elektronischen OSCE-Checklisten vermieden werden konnten.

Schlussfolgerungen: Elektronische OSCE-Checklisten wurden von einer breiten nationalen Stichprobe von Prüfenden als deutlich benutzerfreundlicher eingestuft und gegenüber OSCE-Checklisten auf Papier bevorzugt, was frühere Ergebnisse von Prüfungen auf Fakultätsebene bestätigt. Außerdem wurden mit den elektronischen OSCE-Checklisten fehlende Bewertungen vermieden. Insgesamt ist die Verwendung elektronischer OSCE-Checklisten daher empfehlenswert.

Schlüsselwörter: OSCE, Checklisten, elektronisch, Usability, Evaluation, national

Felicitas L. Wagner¹
Sabine Feller¹
Felix M. Schmitz²
Philippe G.
Zimmermann²
Rabea Krings¹
Sissel Guttormsen³
Sören Huwendiek¹

1 Universität Bern, Institut für Medizinische Lehre, Abteilung für Assessment und Evaluation, Bern, Schweiz

2 Universität Bern, Institut für Medizinische Lehre, Abteilung für Software-Entwicklung, Usability-Consulting und IT-Infrastruktur, Bern, Schweiz

3 Universität Bern, Institut für Medizinische Lehre, Bern, Schweiz

Hintergrund

Die Durchführung von objektiven, strukturierten klinischen Prüfungen (Objective Structured Clinical Examinations, OSCEs) [1] hat sich als Standardmethode zur zuverlässigen Bewertung praktischer Fähigkeiten in der medizinischen Ausbildung etabliert [2], [3], [4], [5]. Bei einem OSCE werden die klinischen und kommunikativen Fähigkeiten der Studierenden geprüft, während sie mehrere Stationen absolvieren, bei denen sie jeweils ein bestimmtes medizinisches Problem bearbeiten müssen. Während eines OSCE wird ihre Leistung von den Prüfenden anhand standardisierter, strukturierter Checklisten beobachtet und bewertet.

Traditionell wurden Checklisten auf Papier verwendet, um die Kandidierenden während eines OSCE zu bewerten, während in jüngerer Zeit, mit dem Aufkommen mobiler elektronischer Geräte, elektronische OSCE-Checklisten entwickelt und erfolgreich eingesetzt wurden [6], [7], [8]. Elektronische OSCE-Checklisten werden in der Regel auf „Tablet“-Computern präsentiert, da diese flexibler sind als Laptops oder Desktop-Geräte [7]. Bislang gibt es jedoch nur wenige Studien, die elektronische und papierbasierte OSCE-Checklisten hinsichtlich ihrer Usability und dem von den Prüfenden bevorzugten Checklisten-Typ verglichen haben. Da die Prüfenden eine wichtige Interessengruppe und die Nutzer der Checklisten sind, ist es wichtig, wie sie die Usability wahrnehmen und welche Präferenzen sie haben. Nach Petersen [9] besteht das zugrundeliegende Konstrukt der Usability darin, „Objekte mit dem Benutzer kompatibel zu machen“. Bei den Objekten kann es sich um Hardware, Software oder alltägliche Hilfsmittel handeln. Ihr Design sollte die Erledigung der Aufgaben, für die sie entwickelt wurden, nicht behindern. Petersen beschreibt die Usability als die Frage, wie einfach ein Objekt zu benutzen ist und unterstreicht die Bedeutung der Qualität der erzeugten Ergebnisse, für die ein Objekt geschaffen wurde, um seine Usability zu bestimmen. OSCEs stellen eine hohe kognitive Belastung [10] für die Prüfenden dar. Die Prüfenden müssen gleichzeitig die Leistungen der Kandidierenden beobachten und bewerten, die entsprechenden Items in der OSCE-Checkliste finden, über die Anzahl der Punkte entscheiden, die die Kandidierenden für die verschiedenen Items der Checkliste erhalten und die OSCE-Checkliste handhaben. Papierbasierte OSCE-Checklisten können beispielsweise mehrere Seiten lang sein [6]. Außerdem sind die Prüfungszeiten in der Regel lang. In Anbetracht dieser Aspekte müssen die OSCE-Checklisten einfach zu handhaben sein und den Prüfenden eine effiziente und effektive Bewertung der Kandidierenden ermöglichen, um die kognitive Belastung nicht zu erhöhen und die Erfüllung der Aufgabe, für die sie erstellt wurden, nämlich die Bewertung der Leistung der Kandidierenden, nicht zu behindern. Es ist daher wichtig, dass sie die Prüfenden nicht von der Beobachtung und Bewertung der Leistung der Kandidierenden während der Prüfung ablenken. Schließlich ist auch die Qualität des Outputs ein wichtiger Aspekt im Zusammenhang mit der Usability [9]. Papierba-

sierte OSCE-Checklisten können spezifische Probleme hinsichtlich der Datenqualität aufweisen. Sie enthalten oft eine beträchtliche Anzahl von fehlenden Bewertungen, da Checklistenitems während einer Prüfung leicht übersehen werden, was eine zeitaufwändige manuelle Datenüberprüfung erforderlich macht [6]. Darüber hinaus müssen papierbasierte Checklisten in der Regel eingescannt werden, um die Ergebnisse in eine digitale Form zu übertragen, was ein notwendiger Schritt ist, um die Ergebnisse für die Analyse verfügbar zu machen. Die Verwendung elektronischer OSCE-Checklisten führt dagegen zu einer besseren Datenqualität, da fehlende Bewertungen vermieden werden können. Außerdem können die Daten direkt im Anschluss an die Prüfung elektronisch heruntergeladen und gespeichert werden, was die Aufbereitung der Daten für die Analyse im Vergleich zu papierbasierten OSCE-Checklisten weniger fehleranfällig und zeitaufwändig macht [6].

In einer systematischen Übersichtsarbeit aus dem Jahr 2019 [11] wurden drei Studien erwähnt, die die Wahrnehmung elektronischer OSCE-Checklisten durch die Prüfenden untersuchten, wobei jedoch nur zwei dieser Studien [6], [12] die wahrgenommene Usability von papierbasierten gegenüber elektronischen OSCE-Checklisten untersuchten. Zusammen mit den Studien von Hochlehnert et al. [7] und Currie et al. [8] deuten diese beiden Studien [6], [12] darauf hin, dass die wahrgenommene Usability der jeweils verwendeten elektronischen OSCE-Checklisten-Tools hoch war. An diesen Studien war jedoch nur eine kleine bis mittlere Anzahl von Prüfenden beteiligt ($n=10$ [6], $n=35$ und 33 [7], $n=93$ [8], $n=43$ [12]), die einer einzelnen Fakultät angehörten, wodurch die Generalisierbarkeit der Ergebnisse eingeschränkt ist. Wie oben beschrieben, sind fehlende Bewertungen ein kritisches Problem bei OSCE-Checklisten auf Papier. Bislang ist jedoch wenig über das Ausmaß dieses Problems bekannt. Unseres Wissens haben sich bisher nur drei Studien [6], [8], [13] mit diesem Problem befasst.

Unsere Forschungsfragen lauteten daher:

1. Wie bewertet eine große Gruppe von Prüfenden aus fünf verschiedenen Fakultäten, die Erfahrung mit papierbasierten und elektronischen OSCE-Checklisten haben, die Usability von elektronischen gegenüber papierbasierten OSCE-Checklisten in einer nationalen Eignungsprüfung?
2. Welcher Checklisten-Typ (elektronisch oder papierbasiert) wird von den Prüfenden einer nationalen Eignungsprüfung bevorzugt?
3. Wie hoch ist die Anzahl der fehlenden Bewertungen bei papierbasierten OSCE-Checklisten?

Methode

Setting

Die Eidgenössische Clinical Skills-Prüfung Humanmedizin in der Schweiz (die sich an den Prinzipien einer OSCE-Prüfung orientiert) umfasste 2014 und 2015 jeweils 12

Zutreffende Felder so ausfüllen: ●
Korrekturen deutlich anbringen: ●

Bitte markieren Sie **jedes** Kriterium!

	Ja	+/-	Nein
Anamnese			
A1 Schmerz: fragt nach Ort, Art, Ausstrahlung, Schwere, Dauer, Beginn, Entwicklung 5-7 = Ja, 4 = +/-	☐	☐	☐
A2 Modifizierende Faktoren: fragt nach entlastenden Faktoren, erschwerenden Faktoren 2 = Ja, 1 = +/-	☐	☐	☐
A3 Fragt nach auslösenden Ereignissen wie Trauma oder schwerem Heben	☐		☐
A4 Fragt nach Verschlimmerung der Schmerzen bei Valsalva-Manöver	☐		☐
A5 Fragt nach assoziierten Symptomen		☐	
a) Kraftverlust, Gefühlsverlust 2 = Ja, 1 = +/-	☐	☐	☐
b) Fieber, Gewichtsverlust 1 = Ja	☐		☐
c) Harninkontinenz, Stuhlinkontinenz 1 = Ja	☐		☐
A6 Erfragt medizinische Vorgeschichte	☐		☐
Körperliche Untersuchung			
S1 Überprüft Gang, Körperhaltung 2 = Ja, 1 = +/-	☐	☐	☐
S2 Palpation im unteren Rückenbereich Ausreichend erledigt / Teilweise erledigt / nicht gemacht	☐	☐	☐
S3 Untersuchung der unteren Gliedmassen			
a) Reflexe Ausreichend erledigt / Teilweise erledigt / nicht gemacht	☐	☐	☐

Abbildung 1: Beispielbild eines Ausschnitts der papierbasierten OSCE-Checkliste (fiktiver Kandidat)

verschiedene OSCE-Stationen. Die Prüfung wurde an drei aufeinanderfolgenden Tagen an fünf Fakultäten für Humanmedizin in Basel, Bern, Genf, Lausanne und Zürich durchgeführt.

Die papierbasierten OSCE-Checklisten, die für die Prüfung 2014 verwendet wurden, wurden mit einem Teleform-Informationserfassungssystem (Cardiff Software) erstellt. Diese Checklisten mussten nach der Prüfung eingescannt werden (DR-6050 C Image Formula Scanner; Canon), um die Ergebnisse für die Analyse zur Verfügung zu stellen. Die papierbasierten OSCE-Checklisten enthielten zwischen 25 und 38 Items, was auf die unterschiedlichen Inhalte der einzelnen Stationen zurückzuführen war. Die Items waren immer nach denselben vier Dimensionen sortiert und zwar in folgender Reihenfolge:

1. Anamnese,
2. Status,
3. Management,
4. Kommunikation.

Die papierbasierten OSCE-Checklisten enthielten außerdem zwei globale Bewertungen für den Gesamteindruck. Am Anfang der papierbasierten OSCE-Checklisten befand sich zudem ein Erinnerungssatz, dass alle Punkte der Checkliste ausgefüllt werden müssen.

Die elektronischen OSCE-Checklisten wurden bei der Prüfung 2015 mit identischen iPad-Geräten der vierten Generation (2014, Apple) ausgefüllt. Für die Bewertung der Kandidierenden wurde die iPad-App „OSCE-Eval“ [<https://eosce.ch/>], Version 2.1, verwendet. Diese App enthält verschiedene Funktionen, um potenzielle Fehler zu reduzieren und sicherzustellen, dass die gesammelten Daten für spätere Analysen leicht verfügbar sind. Die Benutzer erhalten visuelles Feedback um sicherzustellen, dass jedes Item ausgefüllt wird. Jedes Item musste bewertet werden, um die Checkliste endgültig einzureichen. Nach der abschließenden Bewertung waren keine Änderungen mehr möglich. Die verschlüsselten Daten wurden

mit einem sicheren Server synchronisiert und die Daten automatisch in eine elektronische Tabelle exportiert. Je nach den spezifischen Inhalten der Stationen enthielten die elektronischen OSCE-Checklisten zwischen 21 und 42 Items. Die Items waren nach denselben vier Dimensionen geordnet wie in der papierbasierten OSCE-Checkliste. Zusätzlich gab es in jeder Checkliste zwei globale Bewertungen für den Gesamteindruck (siehe Abbildung 1 und Abbildung 2 für Beispiele beider Checklisten-Typen). Die elektronischen Checklisten wurden nach den damaligen Usability-Standards entwickelt und es wurden erhebliche Anstrengungen unternommen, um ihre Akzeptanz während der Entwicklung iterativ zu überprüfen.

Bei der Einführung in die OSCE-Prüfung wurden die Prüfenden sowohl 2014 als auch 2015 kurz in die Verwendung der Checklisten und die Bewertung der Kandidierenden eingewiesen. Zusätzlich hatten die Prüfenden, die mit der elektronischen OSCE-Checkliste arbeiteten (im Jahr 2015), die Möglichkeit, sich vor der Prüfung ein spezielles, 6-minütiges Schulungsvideo anzusehen, um sich mit der Checkliste und der Software vertraut zu machen. Dabei handelte es sich um dasselbe Schulungsvideo, das auch von Schmitz et al. [6] verwendet wurde, welche berichteten, dass das Video verständlich und effektiv war.

Stichprobe

Im Jahr 2014 bewerteten 696 Prüfende die Leistungen der Kandidierenden an den verschiedenen Stationen anhand der papierbasierten OSCE-Checkliste. Im Jahr 2015 bewerteten 696 Prüfende die Leistungen der Kandidierenden an den verschiedenen Stationen anhand der elektronischen OSCE-Checkliste. Alle Prüfenden wurden eingeladen, an der Umfrage teilzunehmen. Es wurden nur Daten von Prüfenden berücksichtigt, die mindestens einmal mit beiden Checklisten-Typen (papier-

A Anamnese

- Schmerz: fragt nach
 - Ort
 - Art
 - Ausstrahlung
 - Schwere
 - Dauer
 - Beginn
- Modifizierende Faktoren: fragt nach
 - entlastenden Faktoren
 - erschwerenden Faktoren
- Frägt nach auslösenden Ereignissen wie Trauma oder schwerem Heben
- Frägt nach Verschlimmerung der Schmerzen bei Valsalva-Manöver
- Frägt nach assoziierten Symptomen
 - Kraftverlust
 - Gefühlverlust
 - Fieber
 - Gewichtsverlust
 - Harninkontinenz
 - Stuhlinkontinenz
- Erfragt medizinische Vorgeschichte

B Körperliche Untersuchung

- Überprüft
 - Gang
 - Körperhaltung
- Palpation im unteren Rückenbereich
- Untersuchung der unteren Gliedmassen
 - Reflexe

Abbildung 2: Beispielbild eines Ausschnitts der elektronischen OSCE-Checkliste (fiktiver Kandidat)

basiert und elektronisch) gearbeitet hatten, um einen Vergleich zwischen der elektronischen und der papierbasierten OSCE-Checkliste zu ermöglichen.

Material

Die Prüfenden bewerteten den OSCE-Checklisten-Typ (d. h. papierbasiert im Jahr 2014, elektronisch im Jahr 2015), den sie zur Bewertung der Kandidierenden verwendet hatten, anhand eines Fragebogens (siehe Tabelle 1). Der Fragebogen enthielt insgesamt 14 geschlossene Fragen: demographische Angaben (drei Fragen), Erfahrung mit Touchscreen-Geräten und den beiden verschiedenen OSCE-Checklisten-Typen (elektronisch, papierbasiert; drei Fragen), subjektive Usability (sieben Fragen) und zukünftige Präferenz für den Checklisten-Typ (eine Frage). Die Antworten basierten auf einer 7-stufigen Likert-Skala, die von 1 (trifft gar nicht zu) bis 7 (trifft voll zu) reichte.

Die Fragen zur Usability orientierten sich an der System Usability Scale [14]. Der Fragebogen enthielt jedoch nur sieben Fragen zur Usability, um den Aufwand für die Prüfenden zu reduzieren und sechs weitere Fragen, die sich auf die spezifischen Anforderungen in einer OSCE-Prüfungssituation beziehen (z. B. Fragen zur Geschwindigkeit der Bewertung der Kandidierenden, Ablenkung durch die Beobachtung der Kandidierenden, Geschwindigkeit der Korrektur von Eingabefehlern). Diese Fragen spiegeln zudem Usability-Aspekte im Sinne von Petersen [9] wider (Benutzerfreundlichkeit, keine Behinderung der Aufgabenerfüllung).

Der Fragebogen lag in deutscher und französischer Sprache vor. Um die Verständlichkeit der Fragen zu gewährleisten, wurden in beiden Sprachen Think-Alouds [15] mit drei Prüfenden durchgeführt. Bei jedem Think-Aloud wurden die Teilnehmenden gebeten, ihre Gedanken zu verbalisieren, während sie den Fragebogen bearbeiteten. Auf diese Weise konnten Unklarheiten und Missverständnisse aufgedeckt werden und es wurde sichergestellt, dass die Teilnehmenden die Fragen richtig verstanden. Der gleiche Fragebogen wurde 2014 und 2015 verwendet.

Um schließlich die Vollständigkeit der Daten zu ermitteln, wurde die Anzahl der fehlenden Bewertungen bei den papierbasierten Checklisten erfasst.

Vorgehen

In beiden Jahren wurden alle Prüfenden innerhalb der ersten Woche nach der Prüfung eingeladen, den OSCE-Checklisten-Typ zu bewerten, mit dem sie während der Prüfung gearbeitet hatten (d. h. papierbasiert im Jahr 2014, elektronisch im Jahr 2015). Somit war ein Vergleich zwischen den beiden OSCE-Checklisten-Typen möglich. Im Jahr 2014 erhielten die Prüfenden den Fragebogen in Papierform direkt nach der Prüfung, im Jahr 2015 wurde der Fragebogen online mit dem Online-Umfragetool „Unipark“ [https://www.unipark.com/] umgesetzt und die Prüfenden wurden per E-Mail eingeladen. Diese Art von Studie war gemäss den Bestimmungen der Kantonalen Ethikkommission des Kantons Bern, welche mit der Medizinischen Fakultät der Universität Bern assoziiert

Tabelle 1: Evaluationsfragebogen

Usability		
1	Die Checkliste war einfach zu benutzen.	trifft gar nicht zu (1) – trifft voll zu (7)*
2	Die Gliederung der Checkliste war übersichtlich.	trifft gar nicht zu (1) – trifft voll zu (7)*
	Die Checkliste ermöglichte eine ...	
3	rasche Bewertung der Kandidatenleistung	trifft gar nicht zu (1) – trifft voll zu (7)*
4	rasche Erfassung individueller Kommentare zur Kandidatenleistung	trifft gar nicht zu (1) – trifft voll zu (7)*
5	einfache Behebung fehlerhafter Eingaben	trifft gar nicht zu (1) – trifft voll zu (7)*
6	Das Ausfüllen der Checkliste hat mich vom Beobachten des/der Kandidat/in abgelenkt.	trifft gar nicht zu (1) – trifft voll zu (7)*
7	Insgesamt bin ich mit der Checkliste zufrieden.	trifft gar nicht zu (1) – trifft voll zu (7)*
Demographie		
8	Fakultät	Basel / Bern / Genf / Lausanne / Zürich
9	Geschlecht	männlich / weiblich
10	Alter	<35 Jahre / 36-40 Jahre / 41-50 Jahre / 51-60 Jahre / >60 Jahre
11	Wie oft haben Sie bisher (inkl. heute) an OSCEs Kandidierende beurteilt anhand papierbasierter Checklisten?	nie / einmal / 2 mal / 3 mal / 4 mal / >4 mal
12	Wie oft haben Sie bisher (inkl. heute) an OSCEs Kandidierende beurteilt anhand elektronischer Checklisten?	nie / einmal / 2 mal / 3 mal / 4 mal / >4 mal
13	Wie vertraut sind Sie im Umgang mit Touch-Eingabegeräten (Tablets, Smartphones, etc.)?	überhaupt nicht vertraut (1) – sehr vertraut (7)
Präferenz Checklisten-Typ		
14	Falls Sie bereits mit beiden Eingabeformaten gearbeitet haben: mit welchem Typ möchten Sie künftig lieber arbeiten?	papierbasierte Checkliste / elektronische Checkliste / keine Präferenz

Anmerkung: *Nur die Extrempunkte der Skala waren beschriftet.

ist, von einer formellen ethischen Genehmigung ausgenommen. Die Teilnahme an der Umfrage war freiwillig und anonym und die Teilnehmenden wurden durch die Teilnahme an dieser Studie keinen denkbaren Risiken ausgesetzt.

Statistische Analyse

Die Häufigkeiten wurden analysiert um die demographischen Daten der Stichprobe darzustellen. Für Vergleiche zwischen der papierbasierten und der elektronischen OSCE-Checkliste wurden Mann-Whitney-Tests verwendet, da die Daten ordinalskaliert sind. Das Alpha-Niveau wurde gemäß Bonferroni für multiples Testen angepasst. Die Effektgröße wird durch Pearson's r angegeben, die Effektstärke wurde nach Cohen [16] beurteilt.

Ergebnisse

Teilnehmende

Im Jahr 2014 bewerteten 696 Prüfende die Leistungen der Kandidierenden an den verschiedenen Stationen anhand der papierbasierten OSCE-Checkliste. Davon füllten 540 (78%) den Fragebogen aus. Im Jahr 2015 bewerteten 696 Prüfende die Leistungen der Kandidierenden an den verschiedenen Stationen anhand der elektronischen OSCE-Checkliste. Hier lag die Rücklaufquo-

te bei 58%, wobei 406 der 696 Prüfenden den Fragebogen ausfüllten.

Die Fragebogendatensätze von 35 (4%) Prüfenden wurden aufgrund fehlender Daten aus der Datenbank entfernt (2014: 19 Datensätze; 2015: 16 Datensätze). Alle Datensätze, die entfernt wurden, enthielten keine nützlichen Informationen, da diese Personen das Ausfüllen des Fragebogens nach der Sprachauswahl gleich zu Beginn des Fragebogens abbrachen. Die Datensätze von 911 (96%) Prüfenden wurden für die Analyse übernommen. Von diesen 911 Prüfenden gaben 377 (41%) an, dass sie mindestens einmal mit beiden Checklisten-Typen (papierbasiert und elektronisch) gearbeitet hatten. Darunter befanden sich Prüfende aus beiden Jahren, da Prüfende, die 2014 mit der papierbasierten OSCE-Checkliste arbeiteten, bereits bei anderen Prüfungen Erfahrungen mit elektronischen OSCE-Checklisten gemacht haben konnten. Nur die Antworten derjenigen Prüfenden, die Erfahrung mit beiden Checklisten-Typen hatten, wurden in die quantitative Analyse einbezogen. In Tabelle 2 sind die Erfahrungen der Prüfenden mit beiden Checklisten-Typen angegeben.

Demographische Fragen

Die demographischen Daten der Prüfenden der endgültigen Stichprobe sind in Tabelle 2 dargestellt.

Tabelle 2: Demographische Angaben der endgültigen Stichprobe (Prüfende mit Erfahrung sowohl mit der elektronischen als auch mit der papierbasierten OSCE-Checkliste)

Demographie	Detail	Prüfende [n (%)]		
		2014	2015	Total
Fakultät	Basel	32 (28.1)	60 (22.8)	92 (24.4)
	Bern	1 (0.9)	54 (20.5)	55 (14.6)
	Genf	30 (26.3)	49 (18.6)	79 (21.0)
	Lausanne	27 (23.7)	30 (11.4)	57 (15.1)
	Zürich	24 (21.1)	70 (26.6)	94 (24.9)
Geschlecht	männlich	79 (69.3)	182 (69.2)	261 (69.2)
	weiblich	33 (28.9)	81 (30.8)	114 (30.2)
	fehlend	2 (1.8)	0 (0.0)	2 (0.5)
Altersgruppe	<35 Jahre	15 (13.2)	4 (1.5)	19 (5.0)
	35-40 Jahre	17 (14.9)	50 (19.0)	67 (17.8)
	41-50 Jahre	41 (36.0)	99 (37.6)	140 (37.1)
	51-60 Jahre	29 (25.4)	85 (32.3)	114 (30.2)
	>60 Jahre	12 (10.5)	24 (9.1)	36 (9.5)
	fehlend	0 (0.0)	1 (0.4)	1 (0.3)
OSCE Checklisten-Typ				
papierbasiert	einmal	17 (14.9)	48 (18.3)	65 (17.2)
	2 mal	21 (18.4)	53 (20.2)	74 (19.6)
	3 mal	25 (21.9)	50 (19.0)	75 (19.9)
	4 mal	12 (10.5)	33 (12.5)	45 (11.9)
	>4 mal	39 (34.2)	79 (30.0)	118 (31.3)
elektronisch	einmal	73 (64.0)	145 (55.1)	218 (57.8)
	2 mal	22 (19.3)	64 (24.3)	86 (22.8)
	3 mal	9 (7.9)	23 (8.7)	32 (8.5)
	4 mal	5 (4.4)	11 (4.2)	16 (4.2)
	>4 mal	5 (4.4)	20 (7.6)	25 (6.6)

Touch-Screen-Erfahrung

Die meisten Prüfenden aus beiden Jahren hatten große ($M_{2014}=6.21$, $M_{2015}=6.08$; Skala: 1=„überhaupt nicht vertraut“, bis 7=„sehr vertraut“) und vergleichbare Erfahrung mit der Verwendung von Touchscreens ($U=13933.5$, n.s.). Daher ist es unwahrscheinlich, dass die Bewertungen der Prüfenden und die Präferenzen für den Checklisten-Typ (d. h. papierbasiert vs. elektronisch) durch ihre vorherige Erfahrung mit Touchscreen-Geräten beeinflusst wurden.

Usability

Mann-Whitney-Tests ergaben, dass sich die meisten Bewertungen der Usability durch die Prüfenden zwischen den elektronischen und den papierbasierten OSCE-Checklisten signifikant unterschieden (siehe Tabelle 3). Das Bonferroni-korrigierte Alpha-Niveau für die sieben im Folgenden beschriebenen Vergleiche betrug .007. Die elektronische OSCE-Checkliste wurde als signifikant einfacher zu benutzen ($U=9799.5$, $p<0.007$, $r=.33$) und übersichtlicher ($U=11441.5$, $p<0.007$, $r=.21$) als die papierbasierte OSCE-Checkliste bewertet. Die elektronische OSCE-Checkliste schnitt auch bei der Geschwindigkeit

der Bewertung der Kandidierenden ($U=11408$, $p<0.007$, $r=.20$) und der Korrektur von Eingabefehlern ($U=7766.5$, $p<0.007$, $r=.41$) signifikant besser ab. Außerdem wurde das Ausfüllen der elektronischen OSCE-Checkliste als signifikant weniger ablenkend empfunden als das Ausfüllen der papierbasierten OSCE-Checkliste ($U=11673$, $p<0.007$, $r=.18$). Die Geschwindigkeit, mit der Kommentare geschrieben werden konnten, wurde zwischen den beiden Checklisten-Typen nicht als unterschiedlich bewertet ($U=12554$, $p=0.035$, n.s., $r=.11$). Auch die Gesamtzufriedenheit der Prüfenden war bei der elektronischen OSCE-Checkliste signifikant höher als bei der papierbasierten ($U=7499$, $p<0.007$, $r=.42$) (siehe Tabelle 4). Die Effektstärken nach Cohen [16] waren entweder mittel (zwischen 0.3 und 0.5) oder gering (zwischen 0.1 und 0.3). Insgesamt erhielt die elektronische OSCE-Checkliste höhere Usability-Bewertungen als die papierbasierte OSCE-Checkliste.

Checklisten-Typ-Präferenz

Am Ende des Fragebogens wurden die Prüfenden aus den Jahren 2014 und 2015, die Erfahrungen sowohl mit elektronischen als auch mit papierbasierten OSCE-Checklisten angegeben hatten, gefragt, ob sie in Zukunft

Tabelle 3: Usability-Bewertungen der elektronischen und der papierbasierten OSCE-Checkliste

Usability Faktor	Papierbasierte Checkliste						Elektronische Checkliste						p
	n	M	SD	Mdn	Min	Max	n	M	SD	Mdn	Min	Max	
Einfache Benutzung	114	6.13	1.02	6	2	7	262	6.69	0.61	7	4	7	*
Übersichtlichkeit der Checkliste	113	6.16	0.96	6	2	7	263	6.50	0.83	7	3	7	*
Rasche Bewertung	113	5.97	1.00	6	2	7	263	6.32	0.94	7	2	7	*
Rasche Kommentarerfassung	112	5.62	1.32	6	1	7	259	5.19	1.63	5	1	7	n.s.
Einfache Fehlerkorrektur	111	5.36	1.62	6	1	7	260	6.55	0.86	7	2	7	*
Ablenkung von den Kandidierenden ¹	114	3.06	1.85	2	1	7	261	2.48	1.81	2	1	7	*
Zufriedenheit insgesamt	113	5.18	1.60	6	1	7	260	6.39	0.93	7	1	7	*

Anmerkungen: n=Anzahl Antworten, M=Mittelwert, SD=Standardabweichung, Mdn=Median, Min=Minimum, Max=Maximum; Skala von 1="trifft gar nicht zu" bis 7="trifft voll zu"; n.s. = nicht signifikant, * $p < 0.007$

¹ tiefe Werte stehen für eine bessere Bewertung.

Tabelle 4: Zukünftige Präferenz für die Verwendung des OSCE-Checklisten-Typs

Checklisten-Typ	Prüfende [n (%)]	
	2014	2015
Papierbasierte Checkliste	20 (18.2)	3 (1.2)
Elektronische Checkliste	60 (54.5)	230 (89.8)
Keine Präferenz	30 (27.3)	23 (9.0)

Tabelle 5: Anzahl der fehlenden Bewertungen bei den papierbasierten OSCE-Checklisten

Anzahl fehlende Bewertungen	Checklisten	
	n	%
0	8612	85.8
1	1067	10.6
2	230	2.3
3	59	0.6
4	30	0.3
5	29	0.3
6	4	0.04
7	4	0.04
8	2	0.02
9	1	0.01
10	1	0.01
12	1	0.01

lieber mit elektronischen oder papierbasierten OSCE-Checklisten arbeiten würden (siehe Tabelle 4). In beiden Jahren bevorzugte die Mehrheit dieser Prüfenden, die Erfahrungen sowohl mit elektronischen als auch mit papierbasierten OSCE-Checklisten gemacht hatten, das elektronische Format.

Fehlende Bewertungen in der papierbasierten OSCE-Checkliste

Die Anzahl der fehlenden Bewertungen in den papierbasierten Checklisten wurde erfasst (siehe Tabelle 5). In 1428 (14.2%) aller papierbasierten Checklisten wurden Bewertungen ausgelassen. Diese Checklisten enthielten zwischen 1 und 12 fehlende Bewertungen ($M=1.45$, $SD=1.03$). In der Mehrheit dieser Checklisten wurde eine Bewertung ausgelassen.

Diskussion

In dieser Studie, die eine große Stichprobe von Prüfenden mit Erfahrung mit elektronischen und papierbasierten OSCE-Checklisten aus fünf verschiedenen Fakultäten im Rahmen der Eidgenössische Clinical Skills-Prüfung Humanmedizin in der Schweiz über zwei aufeinanderfolgende Jahre umfasste, war die wahrgenommene Usability bei der elektronischen OSCE-Checkliste in allen Aspekten signifikant höher als bei der papierbasierten, mit Ausnahme der Geschwindigkeit beim Verfassen von Kommentaren, bei welcher kein signifikanter Unterschied festgestellt wurde. Außerdem bevorzugte es die überwiegende Mehrheit der Prüfenden, in Zukunft mit der elektronischen OSCE-Checkliste zu arbeiten. Schließlich gab es in 14.2% aller papierbasierten Checklisten fehlende Bewertungen, während es in den elektronischen Checklisten keine gab

(in der elektronischen Checkliste mussten alle Items ausgefüllt werden).

Wir konnten in dieser Studie in einem großen Datensatz von Prüfenden aus fünf verschiedenen Fakultäten der Eidgenössischen Clinical Skills-Prüfung zeigen, dass elektronische OSCE-Checklisten eine höhere wahrgenommene Usability aufweisen, als papierbasierte Checklisten. Bislang wurde die Usability elektronischer OSCE-Checklisten nur in Studien mit einer meist kleinen bis mittleren Anzahl von Prüfenden untersucht [6], [7], [8], [12]. Angesichts der hohen kognitiven Belastung [10], der die Prüfenden während eines OSCE ausgesetzt sind, ist die Usability des OSCE-Checklisten-Typs zur Bewertung der Kandidierenden von wesentlicher Bedeutung. Die Prüfenden in dieser Studie waren insgesamt zufriedener mit der elektronischen OSCE-Checkliste als mit der papierbasierten. Sie empfanden die elektronische OSCE-Checkliste im Vergleich zur papierbasierten OSCE-Checkliste als übersichtlicher und einfacher zu benutzen und fühlten sich während der Prüfung auch weniger abgelenkt. Darüber hinaus wurde die Bewertung der Kandidierenden sowie die Korrektur von Eingabefehlern bei der elektronischen OSCE-Checkliste als einfacher empfunden als bei der papierbasierten OSCE-Checkliste. Es gab jedoch keinen signifikanten Unterschied hinsichtlich der wahrgenommenen Geschwindigkeit beim Schreiben von Kommentaren auf einem Touch-Pad (d. h. der elektronischen OSCE-Checkliste) und beim Schreiben von Kommentaren auf Papier. Es ist davon auszugehen, dass all diese Faktoren dazu beitragen, dass die Prüfenden ihre Aufmerksamkeit besser auf die Kandidierenden und ihre Leistungen richten können während der Prüfung, wodurch die kognitive Belastung der Prüfenden bei der elektronischen OSCE-Checkliste geringer bleibt.

Unsere Analyse der fehlenden Bewertungen in den papierbasierten OSCE-Checklisten ergab, dass in einigen der einzelnen Checklisten tatsächlich eine größere Anzahl von Items fehlte. Insgesamt war die Anzahl der ausgelassenen Items jedoch moderat, da dies nur bei 14.2% aller papierbasierten OSCE-Checklisten der Fall war. In der Mehrzahl dieser Checklisten (10.6% aller Checklisten mit fehlenden Items) wurde nur ein fehlendes Item gefunden. Die durchschnittliche Anzahl der fehlenden Bewertungen in diesen Checklisten betrug insgesamt 1.45, was im Vergleich zu den von Schmitz et al. [6] gemeldeten fehlenden Bewertungen viel geringer ist. Die Anzahl der fehlenden Bewertungen bei papierbasierten OSCE-Checklisten scheint also zwischen verschiedenen Prüfungen und Prüfenden stark zu variieren. Bei der Verwendung elektronischer OSCE-Checklisten hingegen können fehlende Bewertungen gänzlich vermieden werden, was einen wichtigen Vorteil der elektronischen OSCE-Checkliste darstellt, da sie dadurch eine höhere Datenqualität gewährleistet [6], [8], [12]. Die Qualität des Outputs als wichtiger Aspekt der Usability [9] ist somit bei elektronischen OSCE-Checklisten deutlich höher. Trotz der Einführung in die papierbasierte OSCE-Checkliste zu Beginn der Prüfung und eines Erinnerungssatzes oben auf der Checkliste gab es immer noch Fälle, in denen Prüfende

es versäumten, Items zu bewerten. Eine mögliche Erklärung dafür könnte sein, dass aufgrund der geringeren Usability der papierbasierten Checkliste die kognitive Belastung für die Prüfenden in einigen Situationen zu hoch wurde, was zu den versäumten Bewertungen geführt haben könnte.

Die Analyse der Präferenzen für den jeweiligen Checklisten-Typ zeigte ein klares Ergebnis. Sowohl 2014 (54.5%) als auch 2015 (89.9%) gab die Mehrheit der mit beiden OSCE-Checklisten-Typen erfahrenen Prüfenden an, dass sie in Zukunft lieber mit elektronischen OSCE-Checklisten arbeiten würde, was die Ergebnisse von Schmitz et al. [6] und Currie et al. [8] bestätigt. Wir gehen davon aus, dass dieses Ergebnis mit der höheren wahrgenommenen Usability der elektronischen OSCE-Checkliste zusammenhängt.

Die Stärken der vorliegenden Studie sind die große Anzahl der beteiligten Prüfenden, der hohe Anteil dieser Prüfenden, der mit beiden Arten von OSCE-Checklisten (d. h. papierbasiert und elektronisch) Erfahrung hatte und der multi-institutionelle (nationale) Rahmen, der Prüfende aus fünf verschiedenen Fakultäten aus zwei Sprachregionen umfasste.

Eine Einschränkung dieser Studie besteht darin, dass die Daten – abgesehen von den fehlenden Bewertungen – auf den subjektiven Eindrücken der Prüfenden in Bezug auf die Verwendung der Checklisten beruhen und kein objektives Maß für die Usability darstellen.

Schlussfolgerungen

Die Ergebnisse der vorliegenden Studie zeigen, dass die wahrgenommene Usability elektronischer OSCE-Checklisten im Vergleich zu traditionellen papierbasierten Checklisten signifikant höher war, mit Ausnahme der Geschwindigkeit beim Verfassen von Kommentaren, bei der es keinen statistisch signifikanten Unterschied gab. Die Mehrheit der Prüfenden, die sowohl mit elektronischen als auch mit papierbasierten OSCE-Checklisten Erfahrung hatten, zog es vor, in Zukunft mit den elektronischen OSCE-Checklisten zu arbeiten. Wichtig ist, dass fehlende Bewertungen bei papierbasierten Checklisten ein häufiges Problem sind, während sie bei der Verwendung elektronischer OSCE-Checklisten vollständig vermieden werden können, was eine hohe Datenqualität gewährleistet. Insgesamt zeigen die Ergebnisse der vorliegenden Studie mit einer großen Zahl von Prüfenden aus verschiedenen Fakultäten, dass elektronische OSCE-Checklisten, wenn sie mit Schwerpunkt auf Usability entwickelt werden, Vorteile haben, die über technische Faktoren hinausgehen und daher als empfehlenswerte Alternative zu papierbasierten OSCE-Checklisten betrachtet werden können.

Deklarationen

Ethische Genehmigung und Zustimmung zur Teilnahme

Diese Art von Studie wurde gemäss den Bestimmungen der Kantonalen Ethikkommission der Medizinischen Fakultät der Universität Bern als von einer formellen ethischen Genehmigung ausgenommen betrachtet. Dennoch bestätigen wir, dass diese Studie in Übereinstimmung mit der Deklaration von Helsinki durchgeführt wurde, soweit sie für diese Art von Studie anwendbar ist und dass die Teilnehmer auf freiwilliger Basis teilgenommen haben, dass sie durch das präsentierte Material nicht identifiziert werden können und dass sie durch die Teilnahme an dieser Studie keinen denkbaren Risiken ausgesetzt wurden.

Einverständnis zur Veröffentlichung

Nicht anwendbar. Die Abbildungen in diesem Manuskript zeigen ein fiktives Beispiel, das speziell für diese Veröffentlichung erstellt wurde, um die elektronischen und papierbasierten OSCE-Checklisten zu veranschaulichen.

Verfügbarkeit von Daten und Materialien

Die im Rahmen der aktuellen Studie analysierten Datensätze sind auf begründete Anfrage bei der entsprechenden Autorin erhältlich. Weitere Informationen über die OSCE-Eval-Anwendung finden Sie auf der Website [<https://eosce.ch/>].

Förderung

Diese Studie wurde vom Institut für Medizinische Lehre der Universität Bern, Schweiz, finanziert.

Beiträge der Autoren

F.W. analysierte und interpretierte die Daten und schrieb das Manuskript mit Unterstützung von S.H., R.K. half bei der Analyse der Daten. S.H., S.F., F.S. und P.Z. erstellten den Fragebogen für die Umfrage. S.G., S.F., R.K., F.S., P.Z. und S.H. haben das Manuskript kritisch überarbeitet. Alle Autoren haben das endgültige Manuskript gelesen und genehmigt.

Danksagung

Wir danken: den Prüfenden für die Teilnahme an der Umfrage; den Verantwortlichen vor Ort, Monica Escher, Felix Eymann, Claudia Glauser, David Gachoud und Roger Kropf; den Mitgliedern der Arbeitsgruppe für klinische Fähigkeiten, Gianmarco Balestra, Christoph Berendonk, Sylvie Félix, Sabine Feller, Philippe Huber, Matteo Monti,

Kai Schnabel und Ernst Jünger; und dem IT-Support-Team der Eidgenössischen Zulassungsprüfung an den Fakultäten Basel, Bern, Genf, Lausanne und Zürich. Ferner danken wir Daniel Stricker für seine Beratung bei der Analyse der fehlenden Items. Schliesslich möchten wir dem Bundesamt für Gesundheit und der Prüfungskommission für die Unterstützung dieser Prüfung danken.

Interessenkonflikt

Die Autor*innen erklären, dass sie keinen Interessenkonflikt im Zusammenhang mit diesem Artikel haben.

Die OSCE-Eval-Anwendung, mit der die elektronische OSCE-Checkliste implementiert wurde, wurde von der Abteilung für Softwareentwicklung, Usability-Beratung und IT-Infrastruktur (ASCI) am Institut für Medizinische Ausbildung konzipiert und entwickelt. Weitere Informationen über die Anwendung können über die Website [<https://eosce.ch/>] abgerufen werden.

Literatur

1. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *Br Med J.* 1975;1(5955):447-451. DOI: 10.1136/bmj.1.5955.447
2. Barman A. Critiques on the objective structured clinical examination. *Ann Acad Med Singap.* 2005;34(8):478-482.
3. Reznick RK, Smee S, Baumber JS, Cohen R, Rothman A, Blackmore D, Berard M. Guidelines for estimating the real cost of an objective structured clinical examination. *Acad Med.* 1993;68(7):513-517. DOI: 10.1097/00001888-199307000-00001
4. Smee S. Skill based assessment. *Br Med J.* 2003;326(7391):703-706. DOI: 10.1136/bmj.326.7391.703
5. Zayyan M. Objective Structured Clinical Examination: The Assessment of Choice. *Oman Med J.* 2011;26(4):219-222. DOI: 10.5001/omj.2011.55
6. Schmitz FM, Zimmermann PG, Gaunt K, Stolze M, Guttormsen Schär S. Electronic Rating of Objective Structured Clinical Examinations: Mobile Digital Forms Beat Paper and Pencil Checklists in a Comparative Study. In: Holzinger A, Simoncic KM, editors. *Information Quality in eHealth.* Berlin: Springer; 2011. p.501-512. DOI: 10.1007/978-3-642-25364-5_35
7. Hochlehnert A, Schultz JH, Möltner A, Timbil S, Brass K, Jünger J. Electronic acquisition of OSCE performance using tablets. *GMS Z Med Ausbild.* 2015;32(4):Doc41. DOI: 10.3205/zma000983
8. Currie GP, Sinha S, Thomson F, Cleland J, Denison AR. Tablet computers in assessing performance in a high stakes exam: opinion matters. *J R Coll Physicians Edinb.* 2017;47(2):164-167. DOI: 10.4997/JRCPE.2017.215
9. Petersen D. Usability theory, practice and evaluation for learning objects. In: Koohang A, Harman K, editors. *Learning objects: Applications, implications, & future directions.* Informing Science; 2007. p.337-370.
10. Young JQ, Van Merriënboer J, During S, Ten Cate O. Cognitive load theory: implications for medical education: AMEE Guide No. 86. *Med Teach.* 2014;36(5):371-384. DOI: 10.3109/0142159X.2014.889290

11. Philips A, Fordyce Mackintosh S, Gibbs C, Ng L, Fryer CE. A comparison of electronic and paper-based clinical skills assessment: Systematic review. *Med Teach*. 2019;41(10):1151-1159. DOI: 10.1080/0142159X.2019.1623387
12. Treadwell I. The usability of personal digital assistants (PDAs) for assessment of practical performance. *Med Educ*. 2006;40(9):855-861. DOI: 10.1111/j.1365-2929.2006.02543.x
13. Denison A, Bate E, Thompson J. Tablet versus paper marking in assessment: feedback matters. *Perspect Med Educ*. 2016;5(2):108-113. DOI: 10.1007/s40037-016-0262-8
14. Brooke J. SUS-A quick and dirty usability scale. In: Jordan PJ, Thomas B, Weerdmeester BA, McClelland IL, editors. *Usability evaluation in industry*. London: Taylor & Francis; 1996. p.4-7.
15. Collins D. Pretesting survey instruments: An overview of cognitive methods. *Qual Life Res*. 2003;12(3):229-238. DOI: 10.1023/A:1023254226592
16. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Erlbaum; 1988.

Korrespondenzadresse:

Felicitas L. Wagner
 Universität Bern, Institut für Medizinische Lehre, Abteilung
 für Assessment und Evaluation, Mittelstr. 43, CH-3012
 Bern, Schweiz
felicitas.wagner@iml.unibe.ch

Bitte zitieren als

Wagner FL, Feller S, Schmitz FM, Zimmermann PG, Krings R, Guttormsen S, Huwendiek S. Usability and preference of electronic vs. paper and pencil OSCE checklists by examiners and influence of checklist type on missed ratings in the Swiss Federal Licensing Exam. *GMS J Med Educ*. 2022;39(2):Doc24. DOI: 10.3205/zma001545, URN: urn:nbn:de:0183-zma0015453

Artikel online frei zugänglich unter

<https://doi.org/10.3205/zma001545>

Eingereicht: 13.01.2021

Überarbeitet: 28.01.2022

Angenommen: 09.02.2022

Veröffentlicht: 14.04.2022

Copyright

©2022 Wagner et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.