

A pilot survey of student perceptions on the benefit of the OSCE and MCQ modalities

Abstract

Objective: The objective structured clinical examination (OSCE) has become widely accepted as a form of assessment in medical education. At the same time, the more traditional multiple choice question (MCQ) examinations remain a central modality of student assessment. This pilot survey aimed to investigate students' perceptions about the benefits of the OSCE and MCQs to yield data supporting the implementation of this assessment strategy into the national medical licensing examination in Germany.

Methods: A questionnaire was delivered electronically to 34 German medical schools. Students in years 3-6 were invited to rate 11 items about objectives of good medical assessment. All items were presented for both the OSCE and MCQs using a 5-point Likert Scale (1=strongly disagree to 5=strongly agree). Factor analysis was used to identify underlying components in the ratings. Average scores of items that belonged to a component were computed.

Results: Data analysis included 1,082 students from 32 medical schools. For the OSCE, factor analysis revealed two components, which were labelled "educational impact" and "development of clinical competence". The average scores of items were 3.37 and 3.55, respectively. For the MCQ modality, also two components emerged. These were labelled "perceived weaknesses of MCQs" and "perceived strengths of MCQs" (consisting of items such as "promotes my theoretical knowledge"). The average scores for these components were 1.85 and 3.62.

Conclusion: The results of this pilot survey indicate that students consider both OSCE and MCQs as useful assessments for the purposes for which they were designed. The assessment strategy thus appears appropriate and it should be used in the national licensing examination.

Keywords: medical assessment, OSCE and MCQ modalities, perceptions, students

Stefan Müller¹
Utz Settmacher¹
Ines Koch²
Uta Dahmen³

1 Jena University Hospital,
Department of General,
Visceral and Vascular
Surgery, Jena, Germany

2 Jena University Hospital,
Department of Gynaecology
and Reproductive Medicine,
Jena, Germany

3 Jena University Hospital,
Department of General,
Visceral and Vascular
Surgery, Experimental
Transplantation Surgery,
Jena, Germany

1. Introduction

Assessment is a core aspect of the medical education process. Besides making judgements about a candidate's competence or performance, assessment influences the curriculum and, more importantly, drives students' learning. For producing good physicians, medical assessment should adapt to clinical practice and be able to cover the necessary spectrum of competencies [1], [2], [3], [4]. Traditional (written) assessment practices, however, focus on testing students' knowledge instead of appraising performance related skills [5]. This may explain why medical school graduates are often not adequately prepared for clinical work [6], [7].

The objective structured clinical examination (OSCE) is a performance-based assessment that was developed to appraise a student's clinical performance. An OSCE typically includes a series of stations where examinees are required to apply their knowledge and skills in simulated settings. At each station, examinees' performance is rated

according to pre-established criteria. The modality has become widely accepted as a form of assessment in many countries [8]. Although OSCEs are now widespread in use, the more traditional multiple choice question (MCQ) examinations continue to play a central role in undergraduate medical education [9], [10].

1.1. Medical education in Germany

All 36 medical schools in Germany (i.e. schools founded before 2012) have a six-year undergraduate programme that is based on the statutory provisions of the German licensing regulations for physicians [https://www.gesetze-im-internet.de/_appro_2002/BJNR240500002.html]. The programmes generally comprise three sections: the preclinical years (years 1 and 2), containing basic medical sciences; the clinical years (years 3, 4, and 5), where students are introduced to the various aspects of clinical medicine; and, finally, the clinical internship year (year 6), in which students participate in full-time clinical rota-

tions [11]. At all medical schools, students are required to take in-house exams to register for the national licensing examination.

In conjunction with the amendment of the licensing regulations for physicians [https://www.gesetze-im-internet.de/_appro_2002/BJNR240500002.html] of 2002, the assessment of performance related skills has gained increasing importance. Meanwhile, 34 of the 36 (94.4%) medical schools use OSCE exams as part of their in-house assessment strategy [12]. The use of MCQ examinations, however, still dominates student assessment at German medical schools (own data).

The national medical licensing examination in Germany is a three-part examination. Part I of the examination after year 2 assesses basic medical science subjects. Part II after year 5 is a written assessment to test a student's clinical knowledge. This part of the examination lasting three days consists of standardised MCQ examinations, which are organised centrally by the Institute for Medical and Pharmaceutical Examination Questions [https://www.impp.de/]. Finally, part III of the examination after year 6 includes a not standardised two-day clinical-practical exam. Recently, it has been shown that this part of the national examination does not sufficiently assess a medical student's clinical competence [13]. As opposed to other countries, like Switzerland, Canada or the United States [14], [15], [16], the OSCE modality is not yet used in the German medical licensing examination. However, within the framework of the "master plan medical education 2020" [https://www.bmbf.de/de/masterplan-medizinstudium-2020-4024.html], it is intended to incorporate the OSCE assessment into these high-stakes licensure examinations.

1.2. Purpose of the survey

In this pilot survey, we aimed to investigate students' perceptions about the benefits of the OSCE and MCQ modalities to yield data that support the incorporation of the OSCE into the German medical licensing examination. We focussed on the OSCE and MCQs, as these tools are commonly used methods of assessment at medical schools in Germany, and are, moreover, the key components of other national licensing examinations, for example the Medical Council of Canada Qualifying Examination (MCCQE) and the United States Medical Licensing Examination (USMLE) [15], [16]. We focussed only on the students' views. The attitudes of the teaching staff were out of scope and thus were not considered in this survey.

2. Methods

2.1. Survey population

According to data from the Federal Statistical Office [Statistisches Bundesamt], more than 85,000 students (87,863) were enrolled at German medical schools in 2014/15. The majority of them were female students

(53,352; 60.7%) [https://www.destatis.de/DE/ZahlenFakten/Indikatoren/LangeReihen/Bildung/Irbil05.html]. With exact information about cohort sizes and assessment schedules at each medical school, we computed the number of students in years 3-6 experienced with the OSCE modality at 34,790.

2.2. Material

An 11-item set was developed. The process of development first involved informal interviews with students from the local medical school. During the interviews, the students were asked to indicate what they expect from good medical assessment in terms of their vocational preparation. The statements obtained were then used to construct the set of items. The set was pretested to ensure that the items were clear and understandable. The complete item set was presented on separate pages for both the OSCE and MCQs, each with the heading "What does ... [the modality] do for you?" within a larger questionnaire.¹ All items were placed on a 5-point Likert Scale (1=strongly disagree, 2=disagree, 3=neither agree nor disagree, 4=agree, and 5=strongly agree). Specifically, the items were:

- A "gives me an understanding of medical care"
- B "demonstrates the practices and principles of medical treatment"
- C "gives me feedback on my performance level"
- D "reveals my strengths in medical practice"
- E "reveals my weaknesses in medical practice"
- F "shows me gaps in my education"
- G "enhances my problem-solving and decision-making abilities"
- H "promotes my theoretical knowledge"
- J "reflects the requirements of the medical profession"
- K "allows me to assess my own ability to work as a medical professional"
- L "helps me with my speciality choice"

Moreover, demographic data on gender, age, academic year, and medical school affiliation were collected at the end of the questionnaire.

2.3. Procedure

The research was performed between February and April 2015. All 34 medical schools using the OSCE modality were solicited to forward a cover letter with the link for accessing the questionnaire to their students in years 3 to 6. Participation in the survey was voluntary and anonymous. No incentives were offered for survey completion. Formal approval from the local research ethics committee was not required.

For ensuring that respondents had personal experience with the OSCE modality, we crosschecked each data set received with the assessment schedule of the relevant medical school. Data were then analysed descriptively. To identify underlying components in the students' ratings, we entered the data into an exploratory factor an-

alysis with oblique rotation for each of the two assessment methods. We summed up the responses to all items loading highly onto a factor and divided the sums by the number of items to calculate average scores.

3. Results

3.1. Participants

In all, 1,189 participants completed the questionnaire. One hundred and seven participants were identified who did not have experience with the OSCE modality, gave the same rating for all the items in the set or did not specify their academic year or medical school affiliation (necessary for crosschecking), so their data were excluded from further analysis. Data of 1,082 participants, which corresponds to 3.1% of the target population, from 32 of the 34 medical schools (from two schools there were no respondents at all) were analysed. Of this total, 747 (69.0%) were female students. The age of participants ranged between 19 and 45 years, with a mean of 25.3 and a median of 25 years, covering the full age range of medical students in Germany. The percentages of students in years 4, 5, and 6 were equal (28.0%, 29.8% or 27.0%). The percentage of students in year 3, however, was smaller (15.2%), because of less experience with the OSCE modality.

3.2. Ratings in the 11-item set

Table 1 displays the ratings in the 11-item set for the OSCE and MCQs. The OSCE was rated positive (means above 3.50) or weakly positive (means between 3.25 and 3.50) for items B (3.81), A (3.75), F (3.53), E (3.36), D (3.31), and C (3.28). The MCQ modality, by contrast, received positive or weakly positive ratings for items H (4.21), C (3.34), and F (3.32).

Table 2 records the item ratings related to the students' demographic characteristics. The ratings for the items did not vary substantially with respect to gender, age or stage of training (academic year) of participants.

3.3. Factor analysis

Table 3 shows the results of the factor analysis including the rotated factor coefficients. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.90 for the OSCE and 0.87 for the MCQs, respectively, suggesting that data were appropriate for factor analysis. Two factors or components with eigenvalues >1 were found for both modalities. For the OSCE, items E, D, F, and C loaded highly (coefficients >0.60) on factor 1 with loadings of 0.95, 0.84, 0.83 or 0.68. The average score of the four items was 3.37. After inspecting the content of the items, we labelled the component "educational impact". Items B and A, as well as item H (mean 3.10) loaded highly on factor 2 (factor loadings of 0.80, 0.74, and 0.77, respectively). The average item score was 3.55. Based on the

content of the items, we labelled this component "development of clinical competence". For the MCQ modality, items D, E, K, J, and G (each with a mean below 2.00) had high loadings on factor 1 (factor coefficients of 0.92, 0.89, 0.78, 0.73, and 0.62, respectively), while items H, C, and F loaded highly on factor 2 (coefficients of 0.85, 0.79 or 0.70). The average scores for the two components were 1.85 and 3.62. The components were labelled "perceived weaknesses of MCQs" and "perceived strengths of MCQs".

4. Discussion

This pilot study investigated the benefits of the OSCE and MCQ modalities perceived by students. In the following paragraphs, we discuss the survey findings classified by themes.

Educational impact is considered an important feature of assessment [17], [18]. An assessment, even when it is used for summative purposes, should have a positive effect on students' future learning by providing feedback to the learners about their strengths and weaknesses. The assessment should also identify areas of weakness in teaching practices or the curriculum so educators can make adjustments. In our study, students rated the items C to F (items that contributed to the component "educational impact") positive or weakly positive for the OSCE. This indicates that the OSCE modality can serve as a tool to improve both students' learning and the curriculum, which is in line with previous research [19], [20]. However, it should be pointed out that the provision of detailed feedback to students is difficult to accomplish in a high-stakes licensing OSCE. Nonetheless, the experience gained from taking the assessment as well as the grades achieved can be used by students to enhance their future learning [21].

A practicing physician should be competent in the domains medical knowledge, clinical skills (e.g. taking history from a patient or performing a physical examination), practical procedures (e.g. establishing intravenous access), patient management, communication, and professional behaviour [22], [23]. Helping students develop competencies in the required domains is a major objective of medical education. The item ratings suggest that students perceive the OSCE modality as useful for developing competencies in the domains other than medical knowledge. This conclusion can be drawn from positive ratings for items A and B, vs. a merely mediocre rating for item H (all items that were assigned to the component "development of clinical competence"). Bearing in mind that the OSCE is a performance measure appraising skills and behaviours that are needed in the clinical workplace [8], the present findings are consistent with the OSCE modality's intended focus.

The greatest perceived strength of the MCQs appears to be in the fostering of the acquisition of knowledge. A strong positive rating for the item "promotes my theoretical knowledge" underpins that. However, our findings

Table 1: Ratings in the 11-item set for the OSCE and MCQs

Item	OSCE	MCQs	<i>p</i>
	Mean (SD)	Mean (SD)	
A gives me an understanding of medical care	3.75 (1.04)	2.28 (1.12)	<0.01
B demonstrates the practices and principles of medical treatment	3.81 (1.08)	2.20 (1.04)	<0.01
C gives me feedback on my performance level	3.28 (1.21)	3.34 (1.13)	0.34
D reveals my strengths in medical practice	3.31 (1.19)	1.78 (0.88)	<0.01
E reveals my weaknesses in medical practice	3.36 (1.15)	1.87 (0.96)	<0.01
F shows me gaps in my education	3.53 (1.14)	3.32 (1.13)	<0.01
G enhances my problem-solving and decision-making abilities	3.12 (1.23)	1.93 (1.00)	<0.01
H promotes my theoretical knowledge	3.10 (1.11)	4.21 (0.83)	<0.01
J reflects the requirements of the medical profession	3.22 (1.18)	1.83 (0.82)	<0.01
K allows me to assess my own ability to work as a medical professional	2.85 (1.19)	1.86 (0.87)	<0.01
L helps me with my speciality choice	2.04 (1.06)	1.85 (1.01)	<0.01

Legend: Items were rated on a 5-point Likert Scale, strongly disagree (1), disagree (2), neither agree nor disagree (3), agree (4), strongly agree (5); the Wilcoxon matched pairs test was used; 1,082 participants.

Table 2: Item ratings related to the students' demographic characteristics

	OSCE / MCQs											
	A	B	C	D	E	F	G	H	J	K	L	<i>n</i>
Gender												
Female	3.82 ^a	3.87 ^a	3.32	3.35	3.38	3.56	3.19 ^a	3.19 ^a	3.23	2.88	2.04	747
Male	3.60 ^b	3.68 ^b	3.19	3.23	3.31	3.45	2.96 ^b	2.88 ^b	3.20	2.79	2.03	335
	2.25	2.23	3.34	1.77	1.86	3.36	1.90	4.21	1.81	1.84	1.86	
	2.35	2.12	3.32	1.80	1.89	3.23	2.02	4.21	1.88	1.90	1.82	
Age, years												
<24	3.80	3.93	3.32	3.40	3.45	3.46	3.14	3.03 ^a	3.29	2.93	2.03	349
24-26	3.75	3.78	3.25	3.28	3.34	3.58	3.15	3.06 ^a	3.17	2.84	2.03	449
>26	3.71	3.75	3.32	3.29	3.32	3.55	3.07	3.27 ^b	3.25	2.81	2.06	272
	2.28	2.18	3.32	1.79	1.91	3.26	1.89	4.22	1.82	1.84	1.88	
	2.27	2.24	3.38	1.78	1.86	3.38	1.96	4.28 ^a	1.81	1.84	1.86	
	2.30	2.17	3.27	1.76	1.83	3.32	1.95	4.09 ^b	1.90	1.91	1.80	
Academic year												
Three	3.81	3.88	3.27	3.42	3.38	3.27 ^a	3.07	3.02	3.38 ^a	3.01	2.11	165
Four	3.79	3.85	3.30	3.36	3.43	3.54	3.16	3.06	3.23	2.92	2.07	303
Five	3.62	3.70	3.17	3.14	3.26	3.58 ^b	3.01	3.05	3.08 ^b	2.73	1.91	322
Six	3.81	3.85	3.40	3.37	3.39	3.61 ^b	3.22	3.22	3.28	2.83	2.10	292
	2.12	1.92 ^a	3.16	1.81	1.91	3.27	1.89	4.21	1.81	1.82	1.79	
	2.30	2.28 ^b	3.34	1.81	1.92	3.26	1.92	4.18	1.92	1.93	1.88	
	2.35	2.29 ^b	3.40	1.78	1.86	3.39	1.98	4.20	1.83	1.85	1.91	
	2.28	2.16	3.35	1.72	1.80	3.34	1.92	4.25	1.75	1.80	1.77	

Legend: Data are presented as means; 5-point Likert Scale, strongly disagree (1), disagree (2), neither agree nor disagree (3), agree (4), strongly agree (5); a, b, differences between groups were significant, $p < 0.05$ (independent *t*-test for gender, one-way ANOVA with Bonferroni test for age groups and academic year); *n*, number of participants.

indicate that students perceive the MCQ modality to be only suitable to measure lower order cognitive skills (e.g. factual knowledge). This raises the issue of whether medical schools in Germany too often rely on context-free MCQs, which typically consist of discrete questions that aim to test factual knowledge. If so, and in order to include higher order cognitive skills (e.g. processes of problem solving or decision making), it would be advisable to expand the use of context-rich MCQs, in which the

questions are directly related to a clinical case presentation [24], [25].

Furthermore, similar to the OSCE, the MCQ modality seems to enable students to evaluate both their own achievements and the content of the curriculum, which is most likely driven by the assessment results. Weakly positive ratings for the items “gives me feedback on my performance level” and “shows me gaps in my education”; vs. low ratings under “reveals my strengths in medical

Table 3: Rotated factor matrix (two factors) for both the OSCE and MCQs

Item	OSCE		MCQs	
	Factors		Factors	
	1	2	1	2
A gives me an understanding of medical care	0.17	0.74	0.43	0.38
B demonstrates the practices and principles of medical treatment	0.04	0.80	0.42	0.31
C gives me feedback on my performance level	0.68	0.16	0.09	0.79
D reveals my strengths in medical practice	0.84	0.08	0.92	-0.14
E reveals my weaknesses in medical practice	0.95	-0.11	0.89	-0.14
F shows me gaps in my education	0.83	-0.15	0.09	0.70
G enhances my problem-solving and decision-making abilities	0.51	0.36	0.62	0.13
H promotes my theoretical knowledge	-0.14	0.77	-0.13	0.85
J reflects the requirements of the medical profession	0.40	0.46	0.73	-0.01
K allows me to assess my own ability to work as a medical professional	0.57	0.31	0.78	0.03
L helps me with my speciality choice	0.30	0.39	0.53	0.18

Legend: Exploratory factor analysis with direct oblimin rotation was performed; two factors with eigenvalues >1 emerged for each assessment modality.

practice” or “reveals my weaknesses in medical practice” suggest this inference.

There are several limitations to this study. First, the study is limited to a relatively small sample representing less than 5% of the target population. This is probably because most medical schools did not inform their students by email, but merely put the cover letter with the link for accessing the questionnaire on their websites. For this reason, a large number of students were likely uninformed about the survey. Second, the study was a pilot survey. The item set presented is not exhaustive and gives an overview only. Despite its limitations, we believe that this survey, the first to investigate the perceptions of students across Germany, provides insight into how and to what extent the OSCE and MCQs are useful assessments. Moreover, our sample was quite representative in terms of gender and age, and we had great diversity regarding the stages of training and medical school affiliations of participants, which make it probable that our results can be generalised to the whole medical student population in Germany.

5. Conclusions

The findings of this pilot survey suggest that students consider both assessment modalities, the OSCE and MCQs, to be valuable tools. In summary, the OSCE may have an impact on the educational process and support the development of skills and behaviours required for clinical practice, while the MCQ modality fosters the acquisition of knowledge. Whilst the employment of assessment programmes including a battery of tests will be the most robust strategy to create a global appraisal of a candidate’s knowledge and skills [2], [3], [26], the use of the OSCE and MCQs appears to be an appropriate assessment strategy. This is further evidence of the need

to incorporate the OSCE into the German medical licensing examination in addition to the existing MCQs.

Notes

¹The questionnaire included items related to various issues. Another manuscript on the students’ learning behaviour when preparing for the OSCE and MCQs has been submitted elsewhere.

Acknowledgements

The authors wish to acknowledge the deans’ offices of the medical schools, all the participating students, and Prof. Dr. Thomas Kessler and Prof. Dr. Rolf Steyer, both Institute of Psychology at Jena University, for providing support in carrying out this research.

Competing interests

The authors declare that they have no competing interests.

References

1. Norcini J, Anderson B, Bollela V, Burch V, Costa MJ, Duvivier R, Galbraith R, Hays R, Kent A, Perrott V, Roberts T. Criteria for good assessment: Consensus statement and recommendations from the Ottawa 2010 Conference. *Med Teach*. 2011;33(3):206-214. DOI: 10.3109/0142159X.2011.551559
2. Schuwirth LWT, Ash J. Principles of assessment. In: Walsh K (Hrsg). *Oxford Textbook of Medical Education*. Oxford: Oxford University Press; 2013. S.409-420. DOI: 10.1093/med/9780199652679.003.0035

3. Shumway JM, Harden RM. AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Med Teach*. 2003;25(6):569-584. DOI: 10.1080/0142159032000151907
4. Sturmberg JP, Farmer L. Educating capable doctors—A portfolio approach. Linking learning and assessment. *Med Teach*. 2009;31(3):e85-89. DOI: 10.1080/01421590802512912
5. Crossley J, Humphris G, Jolly B. Assessing health professionals. *Med Educ*. 2002;36(9):800-804. DOI: 10.1046/j.1365-2923.2002.01294.x
6. Burch VC, Nash RC, Zabow T, Gibbs T, Aubin L, Jacobs B, Hift RJ. A structured assessment of newly qualified medical graduates. *Med Educ*. 2005;39(7):723-731. DOI: 10.1111/j.1365-2929.2005.02192.x
7. Lypson ML, Frohna JG, Gruppen LD, Woolliscroft JO. Assessing residents' competencies at baseline: identifying the gaps. *Acad Med*. 2004;79(6):564-570. DOI: 10.1097/00001888-200406000-00013
8. Humphrey-Murto S, Touchie C, Smee S. Objective structured clinical examinations. In: Walsh K (Hrsg). *Oxford Textbook of Medical Education*. Oxford: Oxford University Press; 2013. S.524-536. DOI: 10.1093/med/9780199652679.003.0045
9. Fowell SL, Maudsley G, Maguire P, Leinster SJ, Bligh J. Student assessment in undergraduate medical education in the United Kingdom, 1998. *Med Educ*. 2000;34 Suppl 1:S1-49. DOI: 10.1046/j.1365-2923.2000.0340s1001.x
10. Mavis BE, Cole BL, Hoppe RB. A survey of student assessment in U.S. medical schools: The balance of breadth versus fidelity. *Teach Learn Med*. 2001;13(2):74-79. DOI: 10.1207/S15328015TLM1302_1
11. Chenot JF. Undergraduate medical education in Germany. *GMS Ger Med Sci*. 2009;7:Doc02. DOI: 10.3205/000061
12. Müller S, Dahmen U, Settmacher U. Application of the Objective Structured Clinical Examination (OSCE) in German medical schools: An inventory. *Gesundheitswesen*. 2016. DOI: 10.1055/s-0042-116435
13. Huber-Lang M, Palmer A, Grab C, Boeckers A, Boeckers TM, Oechsner W. Visions and reality: the idea of competence-oriented assessment for German medical students is not yet realised in licensing examinations. *GMS J Med Educ*. 2017;34(2):Doc25. DOI: 10.3205/zma001102
14. Berendonk C, Schirlo C, Balestra G, Bonvin R, Feller S, Huber P, Jünger E, Monti M, Schnabel K, Beyeler C, Guttormsen S, Huwendiek S. The new final Clinical Skills examination in human medicine in Switzerland: Essential steps of exam development, implementation and evaluation, and central insights from the perspective of the national Working Group. *GMS Z Med Ausbild*. 2015;32(4):Doc40. DOI: 10.3205/zma000982
15. Bordage G, Meguerditchian AN, Tamblyn R. Practice indicators of suboptimal care and avoidable adverse events: a content analysis of a national qualifying examination. *Acad Med*. 2013;88(10):1493-1498. DOI: 10.1097/ACM.0b013e3182a356af
16. Swanson DB, Roberts TE. Trends in national licensing examinations in medicine. *Med Educ*. 2016;50(1):101-114. DOI: 10.1111/medu.12810
17. Shepard LA. The role of assessment in a learning culture. *Educ Res*. 2000;29(7):4-14. DOI: 10.3102/0013189X029007004
18. Van der Vleuten CP. The assessment of professional competence: Developments, research and practical implications. *Adv Health Sci Educ Theory Pract*. 1996;1(1):41-67. DOI: 10.1007/BF00596229
19. König S, Wagner P, Markus PM, Becker H. Anders prüfen – anders studieren: Motivation durch OSCE. *Med Ausbild*. 2002;19(2):73-76.
20. Tervo RC, Dimitrievich E, Trujillo AL, Whittle K, Redinius P, Wellman L. The Objective Structured Clinical Examination (OSCE) in the clinical clerkship: an overview. *S D J Med*. 1997;50(5):153-156.
21. Wood DF. Formative assessment. In: Walsh K (Hrsg). *Oxford Textbook of Medical Education*. Oxford: Oxford University Press; 2013. S.478-488. DOI: 10.1093/med/9780199652679.003.0041
22. Simpson JG, Furnace J, Crosby J, Cumming AD, Evans PA, Friedman Ben David M, Harden RM, Lloyd D, McKenzie H, McLachlan JC, McPhate GF, Percy-Robb IW, MacPherson SG. The Scottish doctor – learning outcomes for the medical undergraduate in Scotland: a foundation for competent and reflective practitioners. *Med Teach*. 2002;24(2):136-143. DOI: 10.1080/01421590220120713
23. Swing SR. The ACGME outcome project: retrospective and prospective. *Med Teach*. 2007;29(7):648-654. DOI: 10.1080/01421590701392903
24. McConnell MM, St-Onge C, Young ME. The benefits of testing for learning on later performance. *Adv Health Sci Educ Theory Pract*. 2015;20(2):305-320. DOI: 10.1007/s10459-014-9529-1
25. Schuwirth LW, van der Vleuten CM. Different written assessment methods: what can be said about their strengths and weaknesses? *Med Educ*. 2004;38(9):974-979. DOI: 10.1111/j.1365-2929.2004.01916.x
26. Van der Vleuten CP, Schuwirth LW. Assessing professional competence: from methods to programmes. *Med Educ*. 2005;39(3):309-317. DOI: 10.1111/j.1365-2929.2005.02094.x

Corresponding author:

Prof. Dr. Uta Dahmen

Jena University Hospital, Department of General, Visceral and Vascular Surgery, Experimental Transplantation Surgery, Drackendorfer Str. 1, D-07747 Jena, Germany, Phone: +49 (0)3641/9-325350, Fax: +49 (0)3641/9-325352
uta.dahmen@med.uni-jena.de

Please cite as

Müller S, Settmacher U, Koch I, Dahmen U. A pilot survey of student perceptions on the benefit of the OSCE and MCQ modalities. *GMS J Med Educ*. 2018;35(4):Doc51. DOI: 10.3205/zma001197, URN: urn:nbn:de:0183-zma0011979

This article is freely available from

<http://www.egms.de/en/journals/zma/2018-35/zma001197.shtml>

Received: 2017-10-26

Revised: 2018-05-24

Accepted: 2018-06-24

Published: 2018-11-15

Copyright

©2018 Müller 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/>.

Eine Piloterhebung über den studentisch wahrgenommenen Nutzen von OSCE und MC-Frageformaten

Zusammenfassung

Zielsetzung: Objective Structured Clinical Examination (OSCE) hat sich als eine Form der Prüfung in der medizinischen Ausbildung durchgesetzt. Gleichwohl stellen die mehr traditionellen Multiple-Choice-Fragen (MC-Fragen) weiterhin ein zentrales Verfahren bei studentischen Prüfungen dar. Das Ziel dieser Piloterhebung war die Untersuchung der Wahrnehmung von Studierenden über den Nutzen von OSCE und MC-Fragen. Es sollten Daten gewonnen werden, die eine Implementierung dieser Prüfungsstrategie in das medizinische Staatsexamen der Bundesrepublik Deutschland unterstützen.

Methodik: Ein Fragebogen wurde elektronisch an 34 medizinische Fakultäten in Deutschland gesandt. Studierende im 3. bis 6. Studienjahr wurden gebeten, 11 Items über Ziele guter medizinischer Prüfungen zu bewerten. Alle Items wurden sowohl für OSCE als auch MC-Fragen auf einer 5-Punkt-Likert-Skala (1=trifft überhaupt nicht zu; 5=trifft vollkommen zu) präsentiert. Eine Faktorenanalyse wurde eingesetzt, um zugrunde liegende Komponenten in den Ratings zu identifizieren. Es wurden Mittelwertindizes aus Items gebildet, die jeweils zu einer Komponente gehörten.

Ergebnisse: Die Datenanalyse umfasste 1.082 Studierende von 32 medizinischen Fakultäten. Für das OSCE-Format ergab die Faktorenanalyse zwei Komponenten, die „Bildungswirkung“ und „Entwicklung klinischer Kompetenz“ benannt wurden. Die Mittelwertindizes der Items lagen bei 3,37 bzw. 3,55. Für das MC-Frageformat ergaben sich ebenfalls zwei Komponenten. Diese wurden „empfundene Schwächen von MC-Fragen“ und „empfundene Stärken von MC-Fragen“ (bestehend aus Items wie „fördert mein theoretisches Wissen“) benannt. Die Mittelwertindizes der Komponenten betrugen 1,85 und 3,62.

Schlussfolgerung: Die Ergebnisse dieser Piloterhebung zeigen, dass die Studierenden OSCE und MC-Fragen als sinnvolle Prüfungen betrachten, und zwar jeweils für die Ziele, für die sie konzipiert wurden. Die Prüfungsstrategie erscheint daher angemessen und sollte im Staatsexamen angewandt werden.

Schlüsselwörter: Medizinische Prüfungen, OSCE und MC-Frageformate, Wahrnehmung, Studierende

Stefan Müller¹
Utz Settmacher¹
Ines Koch²
Uta Dahmen³

1 Universitätsklinikum Jena, Klinik für Allgemein-, Viszeral- und Gefäßchirurgie, Jena, Deutschland

2 Universitätsklinikum Jena, Klinik und Poliklinik für Frauenheilkunde und Fortpflanzungsmedizin, Jena, Deutschland

3 Universitätsklinikum Jena, Klinik für Allgemein-, Viszeral- und Gefäßchirurgie, Experimentelle Transplantationschirurgie, Jena, Deutschland

1. Einführung

Prüfungen sind ein wesentlicher Bestandteil des Ausbildungsprozesses in der Medizin. Neben der Erfassung der Kompetenz oder Performance eines Kandidaten haben Prüfungen Einfluss auf den Lehrplan und steuern darüber hinaus das Lernen der Studierenden. Für das Hervorbringen guter Ärzte sollten sich medizinische Prüfungen an der klinischen Praxis orientieren und in der Lage sein, das dafür notwendige Kompetenzspektrum abzudecken [1], [2], [3], [4]. Klassische (schriftliche) Prüfungspraxis

konzentriert sich jedoch darauf, das Wissen der Studierenden zu testen, anstatt performancebezogene Fähigkeiten zu bewerten [5]. Dies mag erklären, warum Absolventen oftmals nur unzureichend auf die Arbeit in der Klinik vorbereitet sind [6], [7].

Objective Structured Clinical Examination (OSCE) ist eine performancebasierte Prüfung. Das Format wurde entwickelt, um die klinische Performance von Studierenden zu beurteilen. Ein OSCE umfasst in der Regel eine Reihe von Stationen, an denen die Prüflinge ihr Wissen und ihre Fähigkeiten in simulierten Umgebungen anwenden müssen. An jeder Station wird die Performance der Prüflinge nach vorher festgelegten Kriterien bewertet. Das

OSCE-Format hat sich in vielen Ländern als Prüfungsverfahren durchgesetzt [8]. Obgleich OSCE-Prüfungen mittlerweile weit verbreitet sind, spielen die mehr traditionellen Multiple-Choice-Fragen (MC-Fragen) noch immer eine zentrale Rolle in der medizinischen Ausbildung [9], [10].

1.1. Medizinstudium in Deutschland

Alle 36 medizinischen Fakultäten in Deutschland (d. h. Fakultäten, die vor 2012 gegründet wurden) bieten ein sechsjähriges Studium an, das den gesetzlichen Vorgaben der Approbationsordnung für Ärzte (ÄApprO) [https://www.gesetze-im-internet.de/_appro_2002/BJNR240500002.html] entspricht. Das Studium besteht in der Regel aus drei Abschnitten: einem vorklinischen Abschnitt von 2 Jahren, der die medizinischen Grundlagenfächer beinhaltet; einem klinischen Abschnitt von 3 Jahren, in dem den Studierenden die einzelnen klinischen Fächer nähergebracht werden; und abschließend dem Praktischen Jahr (PJ) mit zusammenhängender praktischer Ausbildung in Krankenanstalten [11]. An allen Fakultäten müssen die Studierenden fakultätsinterne Prüfungen ablegen, um sich für die medizinischen Staatsprüfungen anmelden zu können.

In Zusammenhang mit der 2002 durchgeführten Novelisierung der ÄApprO [https://www.gesetze-im-internet.de/_appro_2002/BJNR240500002.html] hat die Überprüfung performancebezogener Fähigkeiten zunehmend an Bedeutung gewonnen. An 34 der 36 (94,4%) Fakultäten sind OSCE-Prüfungen inzwischen Bestandteil der fakultätsinternen Prüfungsstrategie [12]. MC-Aufgaben beherrschen jedoch nach wie vor die studentischen Prüfungen an den Fakultäten (eigene Daten).

Das medizinische Staatsexamen in Deutschland ist eine dreiteilige Prüfung. Der erste Teil der Prüfung nach dem 2. Studienjahr umfasst die medizinischen Grundlagenfächer. Bei dem zweiten Teil nach dem 5. Studienjahr handelt es sich um eine schriftliche Prüfung zur Überprüfung des klinischen Wissens der Studierenden. Dieser Teil der Staatsprüfung, der sich über drei Tage erstreckt, besteht aus standardisierten MC-Prüfungsaufgaben, die zentral vom Institut für medizinische und pharmazeutische Prüfungsfragen (IMPP) [<https://www.impp.de/>] organisiert werden. Der dritte und letzte Teil des Exams nach dem 6. Studienjahr beinhaltet eine nicht standardisierte zweitägige klinisch-praktische Prüfung. Kürzlich wurde gezeigt, dass dieser Teil der staatlichen Prüfung die klinische Kompetenz von Medizinstudierenden nicht zufriedenstellend beurteilt [13]. Im Gegensatz zu anderen Ländern wie der Schweiz, Kanada oder den Vereinigten Staaten [14], [15], [16] wird das OSCE-Format in den deutschen Staatsprüfungen noch nicht eingesetzt. Im Rahmen des „Masterplan Medizinstudium 2020“ [<https://www.bmbf.de/de/masterplan-medizinstudium-2020-4024.html>] ist jedoch vorgesehen, die OSCE-Methode in die Staatsexamensprüfungen mit aufzunehmen.

1.2. Ziel der Erhebung

Diese Piloterhebung hatte zum Ziel, die Wahrnehmung der Studierenden über den Nutzen von OSCE und MC-Fragen zu untersuchen. Es sollten Daten gewonnen werden, die eine Einbeziehung der OSCE-Prüfungsmethode in das medizinische Staatsexamen der Bundesrepublik Deutschland unterstützen. Wir konzentrierten uns auf OSCE und MC-Fragen, da es sich hierbei um gängige Prüfungsverfahren an den medizinischen Fakultäten in Deutschland handelt. Darüber hinaus sind die beiden Formate auch die Hauptbestandteile anderer nationaler Staatsexamina, wie zum Beispiel des Medical Council of Canada Qualifying Examination (MCCQE) und des United States Medical Licensing Examination (USMLE) [15], [16]. Wir haben uns ausschließlich auf die Ansichten der Studierenden konzentriert. Die Einstellungen der Lehrenden lagen außerhalb des Forschungsrahmens und wurden daher in dieser Erhebung nicht berücksichtigt.

2. Methoden

2.1. Befragungspopulation

Nach Zahlen des Statistischen Bundesamtes waren 2014/15 mehr als 85.000 Medizinstudierende (87.863) an den deutschen Fakultäten eingeschrieben. Die Mehrheit davon waren weibliche Studierende (53.352; 60,7%) [<https://www.destatis.de/DE/ZahlenFakten/Indikatoren/LangeReihen/Bildung/Irbil05.html>]. Anhand genauer Informationen zu Kohortengrößen und Prüfungsplänen an den einzelnen Fakultäten errechneten wir die Anzahl der sich zwischen dem 3. und 6. Studienjahr befindenden Studierenden mit OSCE-Erfahrung auf 34.790.

2.2. Erhebungsmaterial

Ein 11-Itemset wurde entwickelt. Der Entwicklungsprozess umfasste zunächst formlose Interviews mit Studierenden der hiesigen medizinischen Fakultät. In den Interviews wurden die Studierenden gebeten, anzugeben, was sie von guten medizinischen Prüfungen im Hinblick auf ihre Berufsvorbereitung erwarten. Die dabei gesammelten Aussagen wurden dann zur Erstellung des Itemsets verwendet. Das Set wurde einem Pretest unterzogen, um sicherzustellen, dass die Items klar und verständlich sind. Das komplette Itemset wurde auf gesonderten Seiten sowohl für OSCE als auch für MC-Fragen innerhalb eines größeren Fragebogens¹ präsentiert und hatte jeweils die Überschrift „Was bringt/nützt Dir ... [das Format]?“. Sämtliche Items wurden auf einer 5-Punkt-Likert-Skala (1=trifft überhaupt nicht zu, 2=trifft kaum zu, 3=weder noch, 4=trifft überwiegend zu, 5=trifft vollkommen zu) abgefragt.

Die Items lauteten im Einzelnen:

- A „gibt mir ein Verständnis medizinischen Vorgehens“
- B „veranschaulicht mir medizinische Behandlungstechniken und Prinzipien“
- C „gibt mir eine Rückmeldung meines Leistungsstandes“
- D „offenbart meine Stärken in ärztlichem Handeln“
- E „offenbart meine Schwächen in ärztlichem Handeln“
- F „zeigt mir Lücken in meiner Ausbildung auf“
- G „stärkt meine Problemlösungs- und Handlungskompetenzen“
- H „fördert mein theoretisches Wissen“
- J „spiegelt die Anforderungen des Arztberufes wider“
- K „erlaubt mir eine Beurteilung meiner ärztlichen Berufsfähigkeit“
- L „hilft mir bei der Wahl meiner späteren Fachrichtung“

Am Ende des Fragebogens wurden noch demografische Angaben zu Geschlecht, Alter, Studienjahr und Fakultätszugehörigkeit erhoben.

2.3. Methodisches Vorgehen

Die Umfrage wurde zwischen Februar und April 2015 durchgeführt. Alle 34 medizinischen Fakultäten, an denen das OSCE-Format eingesetzt wird, wurden gebeten, ihren Studierenden im 3. bis 6. Studienjahr ein Anschreiben mit dem Link zum Fragebogen zukommen zu lassen. Die Teilnahme an der Befragung war freiwillig und anonym. Es wurden keine materiellen Anreize für die Bearbeitung des Fragebogens geboten. Eine formale Zustimmung der zuständigen Ethikkommission war nicht erforderlich.

Um sicherzustellen, dass die Umfrageteilnehmer über persönliche Erfahrungen mit dem OSCE-Format verfügten, haben wir jeden eingegangenen Datensatz mit dem Prüfungsplan der jeweiligen Fakultät abgeglichen. Anschließend wurden die Daten deskriptiv analysiert. Um zugrunde liegende Komponenten in den Ratings der Studierenden zu identifizieren, haben wir die Daten in eine explorative Faktorenanalyse mit schräger Rotation für jede der beiden Prüfungsmethoden eingegeben. Wir haben die Antworten auf alle Items mit jeweils hoher Ladung auf einen Faktor aufsummiert und die Ergebnisse durch die Anzahl der Items geteilt, um Mittelwertindizes zu bilden.

3. Ergebnisse

3.1. Umfrageteilnehmer

Insgesamt haben 1.189 Teilnehmer den Fragebogen ausgefüllt. Es wurden 107 Teilnehmer gefunden, die noch keine Erfahrung mit dem OSCE-Format hatten, alle Items des Sets gleich einstufen oder keine Angaben zu ihrem Studienjahr oder ihrer Fakultätszugehörigkeit machten (notwendig für den Abgleich). Die Daten dieser Teilnehmer wurden von der weiteren Analyse ausgeschlossen. Daten von 1.082 Befragten, was 3,1% der Zielpopulation entspricht, von 32 der 34 medizinischen Fakultäten (von

zwei Fakultäten gab es überhaupt keine Teilnehmer) wurden analysiert. Davon waren 747 (69,0%) weibliche Studierende. Das Alter der Teilnehmer lag zwischen 19 und 45 Jahren, mit einem Mittelwert von 25,3 Jahren und einem Median von 25 Jahren, und deckte damit die gesamte Altersspanne von Medizinstudierenden in Deutschland ab. Die Anteile von Studierenden im 4., 5. und 6. Studienjahr waren in etwa gleich hoch (28,0%, 29,8% bzw. 27,0%), wohingegen der Anteil von Studierenden im 3. Studienjahr aufgrund der geringeren Erfahrung mit dem OSCE-Format niedriger war (15,2%).

3.2. Ratings in dem 11-Itemset

Tabelle 1 zeigt die Ratings in dem 11-Itemset für OSCE und MC-Fragen. Der OSCE bekam positive (Mittelwerte über 3,50) oder schwach positive Ratings (Mittelwerte zwischen 3,25 und 3,50) für die Items B (3,81), A (3,75), F (3,53), E (3,36), D (3,31) und C (3,28). Das MC-Frageformat erhielt dagegen positive oder schwach positive Ratings für die Items H (4,21), C (3,34) und F (3,32).

In Tabelle 2 sind die Itemratings in Abhängigkeit von den demografischen Merkmalen der Studierenden aufgeführt. Die Ratings der Items unterschieden sich nicht wesentlich in Hinblick auf Geschlecht, Alter oder Ausbildungsstand (Studienjahr) der Teilnehmer.

3.3. Faktorenanalyse

Tabelle 3 beinhaltet die Ergebnisse der Faktorenanalyse einschließlich der rotierten Faktorkoeffizienten. Das Kaiser-Meyer-Olkin Maß für die Angemessenheit der Stichprobe lag bei 0,90 für den OSCE bzw. 0,87 für MC-Fragen und zeigte damit, dass die Daten für die Durchführung einer Faktorenanalyse geeignet waren. Es wurden zwei Faktoren bzw. Komponenten mit Eigenwerten >1 für beide Prüfungsformate gefunden. Für den OSCE luden die Items E, D, F und C hoch auf Faktor 1 (Koeffizienten >0,60). Die Items wiesen Ladungen von 0,95, 0,84, 0,83 bzw. 0,68 auf. Der Mittelwertindex aus den vier Items betrug 3,37. Nach einer inhaltlichen Analyse der Items haben wir die Komponente „Bildungswirkung“ benannt. Die Items B und A, als auch Item H (Mittelwert 3,10) hatten hohe Ladungen auf Faktor 2 (Faktorladungen von 0,80, 0,74 bzw. 0,77). Der Mittelwertindex der Items lag bei 3,55. Ausgehend vom Inhalt der Items haben wir diese Komponente „Entwicklung klinischer Kompetenz“ benannt. Beim MC-Frageformat zeigten die Items D, E, K, J und G (Mittelwerte jeweils unter 2,00) hohe Ladungen auf Faktor 1 (Faktorkoeffizienten von 0,92, 0,89, 0,78, 0,73 bzw. 0,62), während die Items H, C und F hoch auf Faktor 2 luden (Koeffizienten von 0,85, 0,79 bzw. 0,70). Die Mittelwertindizes für die beiden Komponenten betrugen 1,85 und 3,62. Die Komponenten wurden „empfundene Schwächen von MC-Fragen“ und „empfundene Stärken von MC-Fragen“ benannt.

Tabelle 1: Ratings in dem 11-Itemset für OSCE und MC-Fragen

Item	OSCE	MC-Fragen	p
	Mittel (SD)	Mittel (SD)	
A gibt mir ein Verständnis medizinischen Vorgehens	3,75 (1,04)	2,28 (1,12)	<0,01
B veranschaulicht mir medizinische Behandlungstechniken und Prinzipien	3,81 (1,08)	2,20 (1,04)	<0,01
C gibt mir eine Rückmeldung meines Leistungsstandes	3,28 (1,21)	3,34 (1,13)	0,34
D offenbart meine Stärken in ärztlichem Handeln	3,31 (1,19)	1,78 (0,88)	<0,01
E offenbart meine Schwächen in ärztlichem Handeln	3,36 (1,15)	1,87 (0,96)	<0,01
F zeigt mir Lücken in meiner Ausbildung auf	3,53 (1,14)	3,32 (1,13)	<0,01
G stärkt meine Problemlösungs- und Handlungskompetenzen	3,12 (1,23)	1,93 (1,00)	<0,01
H fördert mein theoretisches Wissen	3,10 (1,11)	4,21 (0,83)	<0,01
J spiegelt die Anforderungen des Arztberufes wider	3,22 (1,18)	1,83 (0,82)	<0,01
K erlaubt mir eine Beurteilung meiner ärztlichen Berufsfähigkeit	2,85 (1,19)	1,86 (0,87)	<0,01
L hilft mir bei der Wahl meiner späteren Fachrichtung	2,04 (1,06)	1,85 (1,01)	<0,01

Legende: Items wurden auf einer 5-Punkt-Likert-Skala bewertet, trifft überhaupt nicht zu (1), trifft kaum zu (2), weder noch (3), trifft überwiegend zu (4), trifft vollkommen zu (5); der Wilcoxon-Test wurde verwendet; 1.082 Teilnehmer.

Tabelle 2: Itemratings entsprechend den demografischen Merkmalen der Studierenden

	OSCE / MC-Fragen											n
	A	B	C	D	E	F	G	H	J	K	L	
Geschlecht												
Weiblich	3,82 ^a	3,87 ^a	3,32	3,35	3,38	3,56	3,19 ^a	3,19 ^a	3,23	2,88	2,04	747
Männlich	3,60 ^b	3,68 ^b	3,19	3,23	3,31	3,45	2,96 ^b	2,88 ^b	3,20	2,79	2,03	335
	2,25	2,23	3,34	1,77	1,86	3,36	1,90	4,21	1,81	1,84	1,86	
	2,35	2,12	3,32	1,80	1,89	3,23	2,02	4,21	1,88	1,90	1,82	
Alter, Jahre												
<24	3,80	3,93	3,32	3,40	3,45	3,46	3,14	3,03 ^a	3,29	2,93	2,03	349
24-26	3,75	3,78	3,25	3,28	3,34	3,58	3,15	3,06 ^a	3,17	2,84	2,03	449
>26	3,71	3,75	3,32	3,29	3,32	3,55	3,07	3,27 ^b	3,25	2,81	2,06	272
	2,28	2,18	3,32	1,79	1,91	3,26	1,89	4,22	1,82	1,84	1,88	
	2,27	2,24	3,38	1,78	1,86	3,38	1,96	4,28 ^a	1,81	1,84	1,86	
	2,30	2,17	3,27	1,76	1,83	3,32	1,95	4,09 ^b	1,90	1,91	1,80	
Studien-jahr												
Drei	3,81	3,88	3,27	3,42	3,38	3,27 ^a	3,07	3,02	3,38 ^a	3,01	2,11	165
Vier	3,79	3,85	3,30	3,36	3,43	3,54	3,16	3,06	3,23	2,92	2,07	303
Fünf	3,62	3,70	3,17	3,14	3,26	3,58 ^b	3,01	3,05	3,08 ^b	2,73	1,91	322
Sechs	3,81	3,85	3,40	3,37	3,39	3,61 ^b	3,22	3,22	3,28	2,83	2,10	292
	2,12	1,92 ^a	3,16	1,81	1,91	3,27	1,89	4,21	1,81	1,82	1,79	
	2,30	2,28 ^b	3,34	1,81	1,92	3,26	1,92	4,18	1,92	1,93	1,88	
	2,35	2,29 ^b	3,40	1,78	1,86	3,39	1,98	4,20	1,83	1,85	1,91	
	2,28	2,16	3,35	1,72	1,80	3,34	1,92	4,25	1,75	1,80	1,77	

Legende: Daten werden als Mittelwerte angegeben; 5-Punkt-Likert-Skala, trifft überhaupt nicht zu (1), trifft kaum zu (2), weder noch (3), trifft überwiegend zu (4), trifft vollkommen zu (5); a, b, Unterschiede zwischen Gruppen waren signifikant, $p < 0,05$ (t-Test für unabhängige Stichproben bei Geschlecht, Ein-Weg-ANOVA mit Bonferroni-Test bei Altersgruppen und Studienjahr); n, Anzahl der Teilnehmer.

4. Diskussion

Diese Pilotstudie untersuchte den studentisch wahrgenommenen Nutzen von OSCE und MC-Frageformaten. Nachfolgend werden die Befragungsergebnisse nach Themen geordnet diskutiert.

Bildungswirkung („educational impact“) wird als eine wichtige Eigenschaft von Prüfungen gesehen [17], [18]. Eine Prüfung – selbst wenn sie zu summativen Zwecken

verwendet wird – sollte sich positiv auf das weitere Lernen der Studierenden auswirken, indem sie den Lernenden Feedback über deren Stärken und Schwächen gibt. Zugleich sollten Prüfungen Schwachstellen in den Lehrmethoden oder dem Curriculum aufzeigen, sodass die Lehrenden entsprechende Anpassungen vornehmen können. Die Studierenden in unserer Studie bewerteten die Items C bis F (Items, die zur Komponente „Bildungswirkung“ beitragen) positiv oder schwach positiv für den OSCE.

Tabelle 3: Rotierte Faktorenmatrix (zwei Faktoren) für OSCE und MC-Fragen

Item	OSCE		MC-Fragen	
	Faktoren		Faktoren	
	1	2	1	2
A gibt mir ein Verständnis medizinischen Vorgehens	0,17	0,74	0,43	0,38
B veranschaulicht mir medizinische Behandlungstechniken und Prinzipien	0,04	0,80	0,42	0,31
C gibt mir eine Rückmeldung meines Leistungsstandes	0,68	0,16	0,09	0,79
D offenbart meine Stärken in ärztlichem Handeln	0,84	0,08	0,92	-0,14
E offenbart meine Schwächen in ärztlichem Handeln	0,95	-0,11	0,89	-0,14
F zeigt mir Lücken in meiner Ausbildung auf	0,83	-0,15	0,09	0,70
G stärkt meine Problemlösungs- und Handlungskompetenzen	0,51	0,36	0,62	0,13
H fördert mein theoretisches Wissen	-0,14	0,77	-0,13	0,85
J spiegelt die Anforderungen des Arztberufes wider	0,40	0,46	0,73	-0,01
K erlaubt mir eine Beurteilung meiner ärztlichen Berufsfähigkeit	0,57	0,31	0,78	0,03
L hilft mir bei der Wahl meiner späteren Fachrichtung	0,30	0,39	0,53	0,18

Legende: Explorative Faktorenanalyse mit direkter Oblimin-Rotation wurde durchgeführt; es ergaben sich zwei Faktoren mit Eigenwerten >1 für jedes Prüfungsformat.

Das deutet darauf hin, dass das OSCE-Format als ein Hilfsmittel zur Verbesserung des studentischen Lernens und des Lehrplans dienen kann, was in Einklang mit früheren Arbeiten steht [19], [20]. Es sollte jedoch darauf hingewiesen werden, dass die Gabe von detailliertem Feedback in einem Staatsexamen-OSCE nur schwer zu bewerkstelligen ist. Nichtsdestotrotz können die in der Prüfung gemachten Erfahrungen sowie die erreichten Noten von den Studierenden genutzt werden, um ihr weiteres Lernen zu optimieren [21].

Ein praktizierender Arzt sollte in Bezug auf medizinisches Wissen, klinische Fähigkeiten (z. B. Anamneseerhebung oder Durchführung einer körperlichen Untersuchung), praktische Tätigkeiten (z. B. Schaffung eines intravenösen Zugangs), Patientenmanagement, Kommunikation, sowie professionelles Verhalten kompetent sein [22], [23]. Den Studierenden dabei zu helfen, Kompetenzen in den entsprechenden Bereichen zu entwickeln, ist eine Kernaufgabe der medizinischen Ausbildung. Die Itemratings weisen darauf hin, dass das OSCE-Format von den Studierenden als nützlich für die Entwicklung von Kompetenzen in den Bereichen außerhalb von medizinischem Wissen empfunden wird. Die Folgerung ergibt sich aus positiven Ratings für die Items A und B; und einem nur mittelmäßigen Rating für Item H (alles Items, die der Komponente „Entwicklung klinischer Kompetenz“ zugeordnet wurden). Vor dem Hintergrund, dass es sich bei dem OSCE um eine Performanceprüfung zur Beurteilung von Fähigkeiten und Verhaltensweisen handelt, die am klinischen Arbeitsplatz benötigt werden [8], stimmen die vorliegenden Ergebnisse mit der beabsichtigten Schwerpunktsetzung des OSCE-Formats überein.

Die größte von den Studierenden empfundene Stärke der MC-Fragen scheint in der Förderung des Wissenserwerbs zu liegen, was durch ein stark positives Rating für das Item „fördert mein theoretisches Wissen“ untermauert wird. Unsere Ergebnisse zeigen jedoch, dass die Studierenden das MC-Frageformat lediglich für geeignet

halten, kognitive Fähigkeiten niedrigerer Ordnung (z. B. Faktenwissen) zu erfassen. Dies wirft die Frage auf, ob die medizinischen Fakultäten in Deutschland zu häufig auf kontextfreie MC-Fragen zurückgreifen. Diese Form der MC-Prüfung besteht typischerweise aus eigenständigen Fragen, die darauf ausgelegt sind, Faktenwissen abzufragen. Wenn dem so ist, und um kognitive Fähigkeiten höherer Ordnung (z. B. Prozesse der Problemlösung oder Entscheidungsfindung) mit einzuschließen, wäre es für die Fakultäten angebracht, den Einsatz kontextreicher MC-Fragen auszuweiten. Bei solchen MC-Aufgaben werden die Fragen nicht isoliert gestellt, sondern beziehen sich unmittelbar auf eine klinische Falldarstellung [24], [25].

Ähnlich wie bei dem OSCE scheint das MC-Frageformat es den Studierenden zu ermöglichen, sowohl ihre eigenen Leistungen als auch die Unterrichtsinhalte bewerten zu können, was beides höchstwahrscheinlich durch die Prüfungsergebnisse induziert wird. Schwach positive Ratings für die Items „gibt mir eine Rückmeldung meines Leistungsstandes“ und „zeigt mir Lücken in meiner Ausbildung auf“; gegenüber niedrigen Ratings unter „offenbart meine Stärken in ärztlichem Handeln“ oder „offenbart meine Schwächen in ärztlichem Handeln“ legen diesen Schluss nahe.

Diese Studie hat mehrere Einschränkungen. Zum einen ist die Studie auf eine relativ kleine Stichprobe von weniger als 5% der Zielpopulation beschränkt. Dies liegt vermutlich daran, dass die meisten Fakultäten ihre Studierenden nicht per E-Mail benachrichtigt haben, sondern das Anschreiben mit dem Link zum Fragebogen lediglich auf ihre Webseiten gestellt haben. Viele Studierenden waren deshalb wohl nicht über die Umfrage informiert. Zum anderen handelte es sich bei der Studie um eine Piloterhebung. Das vorgestellte Itemset erhebt keinen Anspruch auf Vollständigkeit und gibt nur einen Überblick. Trotz aller Einschränkungen glauben wir, dass diese Erhebung, die als erste die Wahrnehmung der Studierenden

in ganz Deutschland untersuchte, einen Einblick bietet, inwieweit OSCE und MC-Fragen sinnvolle Prüfungen darstellen. Darüber hinaus war unsere Stichprobe in Bezug auf Geschlecht und Alter repräsentativ und wir hatten eine große Vielfalt hinsichtlich des Ausbildungsstandes und der Fakultätszugehörigkeit der Teilnehmer. Das macht es wahrscheinlich, dass unsere Ergebnisse auf die Gesamtpopulation der Medizinstudierenden in Deutschland übertragen werden können.

5. Schlussfolgerung

Die Ergebnisse dieser Piloterhebung zeigen, dass die Studierenden beide Prüfungsformate, OSCE und MC-Fragen, als wertvolle Hilfsmittel betrachten. Zusammenfassend lässt sich sagen, der OSCE hat Auswirkungen auf den Lernprozess und unterstützt die Entwicklung der für die klinische Praxis erforderlichen Fähigkeiten und Verhaltensweisen, während das MC-Frageformat den Wissenserwerb fördert. Auch wenn der Einsatz von Prüfungsprogrammen mit einer Reihe verschiedener Tests wohl die beste Strategie zur Erstellung einer Gesamtbeurteilung der Kenntnisse und Fähigkeiten eines Kandidaten ist [2], [3], [26], stellt die Verwendung von OSCE und MC-Fragen eine geeignete Prüfungsstrategie dar. Die OSCE-Methode sollte daher zusätzlich zu den bereits bestehenden MC-Prüfungsaufgaben in das medizinische Staatsexamen der Bundesrepublik Deutschland aufgenommen werden.

Anmerkung

¹Der Fragebogen enthielt Items zu verschiedenen Themenstellungen. Ein weiteres Manuskript zum Lernverhalten der Studierenden in Vorbereitung auf OSCE und MC-Fragen wurde andernorts eingereicht.

Danksagung

Die Autoren bedanken sich bei den Studiendekanaten der Fakultäten, allen teilnehmenden Studierenden, sowie bei Prof. Dr. Thomas Kessler und Prof. Dr. Rolf Steyer, beide am Institut für Psychologie an der Universität Jena, für die Unterstützung bei der Durchführung dieser Forschungsarbeit.

Interessenkonflikt

Die Autoren erklären, dass sie keine Interessenkonflikte im Zusammenhang mit diesem Artikel haben.

Literatur

1. Norcini J, Anderson B, Bollela V, Burch V, Costa MJ, Duvivier R, Galbraith R, Hays R, Kent A, Perrott V, Roberts T. Criteria for good assessment: Consensus statement and recommendations from the Ottawa 2010 Conference. *Med Teach*. 2011;33(3):206-214. DOI: 10.3109/0142159X.2011.551559
2. Schuwirth LWT, Ash J. Principles of assessment. In: Walsh K (Hrsg). *Oxford Textbook of Medical Education*. Oxford: Oxford University Press; 2013. S.409-420. DOI: 10.1093/med/9780199652679.003.0035
3. Shumway JM, Harden RM. AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Med Teach*. 2003;25(6):569-584. DOI: 10.1080/0142159032000151907
4. Sturmberg JP, Farmer L. Educating capable doctors—A portfolio approach. Linking learning and assessment. *Med Teach*. 2009;31(3):e85-89. DOI: 10.1080/01421590802512912
5. Crossley J, Humphris G, Jolly B. Assessing health professionals. *Med Educ*. 2002;36(9):800-804. DOI: 10.1046/j.1365-2923.2002.01294.x
6. Burch VC, Nash RC, Zabow T, Gibbs T, Aubin L, Jacobs B, Hift RJ. A structured assessment of newly qualified medical graduates. *Med Educ*. 2005;39(7):723-731. DOI: 10.1111/j.1365-2929.2005.02192.x
7. Lyson ML, Frohna JG, Gruppen LD, Woolliscroft JO. Assessing residents' competencies at baseline: identifying the gaps. *Acad Med*. 2004;79(6):564-570. DOI: 10.1097/00001888-200406000-00013
8. Humphrey-Murto S, Touchie C, Smees S. Objective structured clinical examinations. In: Walsh K (Hrsg). *Oxford Textbook of Medical Education*. Oxford: Oxford University Press; 2013. S.524-536. DOI: 10.1093/med/9780199652679.003.0045
9. Fowell SL, Maudsley G, Maguire P, Leinster SJ, Bligh J. Student assessment in undergraduate medical education in the United Kingdom, 1998. *Med Educ*. 2000;34 Suppl 1:S1-49. DOI: 10.1046/j.1365-2923.2000.0340s1001.x
10. Mavis BE, Cole BL, Hoppe RB. A survey of student assessment in U.S. medical schools: The balance of breadth versus fidelity. *Teach Learn Med*. 2001;13(2):74-79. DOI: 10.1207/S15328015TLM1302_1
11. Chenot JF. Undergraduate medical education in Germany. *GMS Ger Med Sci*. 2009;7:Doc02. DOI: 10.3205/000061
12. Müller S, Dahmen U, Settmacher U. Application of the Objective Structured Clinical Examination (OSCE) in German medical schools: An inventory. *Gesundheitswesen*. 2016. DOI: 10.1055/s-0042-116435
13. Huber-Lang M, Palmer A, Grab C, Boeckers A, Boeckers TM, Oechsner W. Visions and reality: the idea of competence-oriented assessment for German medical students is not yet realised in licensing examinations. *GMS J Med Educ*. 2017;34(2):Doc25. DOI: 10.3205/zma001102
14. Berendonk C, Schirlo C, Balestra G, Bonvin R, Feller S, Huber P, Jünger E, Monti M, Schnabel K, Beyeler C, Guttormsen S, Huwendiek S. The new final Clinical Skills examination in human medicine in Switzerland: Essential steps of exam development, implementation and evaluation, and central insights from the perspective of the national Working Group. *GMS Z Med Ausbild*. 2015;32(4):Doc40. DOI: 10.3205/zma000982
15. Bordage G, Meguerditchian AN, Tamblyn R. Practice indicators of suboptimal care and avoidable adverse events: a content analysis of a national qualifying examination. *Acad Med*. 2013;88(10):1493-1498. DOI: 10.1097/ACM.0b013e3182a356af

16. Swanson DB, Roberts TE. Trends in national licensing examinations in medicine. *Med Educ.* 2016;50(1):101-114. DOI: 10.1111/medu.12810
17. Shepard LA. The role of assessment in a learning culture. *Educ Res.* 2000;29(7):4-14. DOI: 10.3102/0013189X029007004
18. Van der Vleuten CP. The assessment of professional competence: Developments, research and practical implications. *Adv Health Sci Educ Theory Pract.* 1996;1(1):41-67. DOI: 10.1007/BF00596229
19. König S, Wagner P, Markus PM, Becker H. Anders prüfen – anders studieren: Motivation durch OSCE. *Med Ausbild.* 2002;19(2):73-76.
20. Tervo RC, Dimitrievich E, Trujillo AL, Whittle K, Redinius P, Wellman L. The Objective Structured Clinical Examination (OSCE) in the clinical clerkship: an overview. *S D J Med.* 1997;50(5):153-156.
21. Wood DF. Formative assessment. In: Walsh K (Hrsg). *Oxford Textbook of Medical Education.* Oxford: Oxford University Press; 2013. S.478-488. DOI: 10.1093/med/9780199652679.003.0041
22. Simpson JG, Furnace J, Crosby J, Cumming AD, Evans PA, Friedman Ben David M, Harden RM, Lloyd D, McKenzie H, McLachlan JC, McPhate GF, Percy-Robb IW, MacPherson SG. The Scottish doctor – learning outcomes for the medical undergraduate in Scotland: a foundation for competent and reflective practitioners. *Med Teach.* 2002;24(2):136-143. DOI: 10.1080/01421590220120713
23. Swing SR. The ACGME outcome project: retrospective and prospective. *Med Teach.* 2007;29(7):648-654. DOI: 10.1080/01421590701392903
24. McConnell MM, St-Onge C, Young ME. The benefits of testing for learning on later performance. *Adv Health Sci Educ Theory Pract.* 2015;20(2):305-320. DOI: 10.1007/s10459-014-9529-1
25. Schuwirth LW, van der Vleuten CM. Different written assessment methods: what can be said about their strengths and weaknesses? *Med Educ.* 2004;38(9):974-979. DOI: 10.1111/j.1365-2929.2004.01916.x
26. Van der Vleuten CP, Schuwirth LW. Assessing professional competence: from methods to programmes. *Med Educ.* 2005;39(3):309-317. DOI: 10.1111/j.1365-2929.2005.02094.x

Korrespondenzadresse:

Prof. Dr. Uta Dahmen

Universitätsklinikum Jena, Klinik für Allgemein-, Viszeral- und Gefäßchirurgie, Experimentelle Transplantationschirurgie, Drackendorfer Str. 1, 07747 Jena, Deutschland, Tel.: +49 (0)3641/9-325350, Fax: +49 (0)3641/9-325352
 uta.dahmen@med.uni-jena.de

Bitte zitieren als

Müller S, Settmacher U, Koch I, Dahmen U. A pilot survey of student perceptions on the benefit of the OSCE and MCQ modalities. *GMS J Med Educ.* 2018;35(4):Doc51.
 DOI: 10.3205/zma001197, URN: urn:nbn:de:0183-zma0011979

Artikel online frei zugänglich unter

<http://www.egms.de/en/journals/zma/2018-35/zma001197.shtml>

Eingereicht: 26.10.2017**Überarbeitet:** 24.05.2018**Angenommen:** 24.06.2018**Veröffentlicht:** 15.11.2018**Copyright**

©2018 Müller 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/>.