

OBJECTIVE STRUCTURED CLINICAL EXAMINATION (OSCE)
IN DER MEDIZINISCHEN AUSBILDUNG DER
BUNDESREPUBLIK DEUTSCHLAND

Dissertation
zur Erlangung des akademischen Grades
doctor medicinae (Dr. med.)

**vorgelegt dem Rat der Medizinischen Fakultät
der Friedrich-Schiller-Universität Jena**

**von Stefan Müller
geboren am 29.11.1982 in Freiburg im Breisgau**

Gutachter

1. Prof. Dr. (med.) Uta Dahmen, Jena
2. Prof. Dr. (med.) Orlando Guntinas-Lichius, Jena
3. Prof. Dr. (med.) Susanne Cupisti, Erlangen

Tag der öffentlichen Verteidigung: 20.10.2020

Inhaltsverzeichnis

Abkürzungsverzeichnis	ii
1 Zusammenfassung	1
2 Einleitung	2
2.1 Prüfungen in der medizinischen Ausbildung	2
2.1.1 Ziele einer Prüfung	2
2.1.2 Prüfungen und Lernen	3
2.2 Überprüfung klinischer Kompetenz	3
2.3 Objective Structured Clinical Examination (OSCE)	4
2.4 Ausbildung und Prüfungspraxis in Deutschland	6
3 Ziele der Arbeit.....	9
4 Publierte Originalarbeiten.....	10
4.1 Objective Structured Clinical Examination (OSCE) an Medizinischen Fakultäten in Deutschland – eine Bestandsaufnahme, S. Müller, U. Dahmen, U. Settmacher, Gesundheitswesen, 80(12):1099–1103, 2018	10
4.2 How the introduction of OSCEs has affected the time students spend studying: results of a nationwide study, Stefan Müller, Ines Koch, Utz Settmacher and Uta Dahmen, BMC Med Educ, 19:146, 2019	11
4.3 A pilot survey of student perceptions on the benefit of the OSCE and MCQ modalities, Stefan Müller, Utz Settmacher, Ines Koch, Uta Dahmen, GMS J Med Educ, 35(4):Doc51, 2018	12
5 Diskussion	13
5.1 Kompetenzorientierte medizinische Ausbildung	14
5.2 Probleme und Herausforderungen	15
5.3 Stärken und Schwächen der Arbeit	16
6 Schlussfolgerungen.....	18
7 Literatur- und Quellenverzeichnis	19
8 Anhang	25

Abkürzungsverzeichnis

ÄApprO	Approbationsordnung für Ärzte
CanMEDS	Canadian Medical Education Directions for Specialists
DOPS	Direct Observation of Procedural Skills
IMPP	Institut für medizinische und pharmazeutische Prüfungsfragen
MC-Prüfung	Multiple-Choice-Prüfung
Mini-CEX	Mini-Clinical Evaluation Exercise
MME	Master of Medical Education
NKLM	Nationaler Kompetenzbasierter Lernzielkatalog Medizin
OSCE	Objective Structured Clinical Examination

1 Zusammenfassung

In der medizinischen Ausbildung sind Prüfungen von zentraler Bedeutung. Sie dienen der Kontrolle des Lernerfolgs, unterstützen den Lernprozess der Studierenden und haben eine lernsteuernde Wirkung. Vor mehr als 40 Jahren wurde die Methode Objective Structured Clinical Examination (OSCE) als ein alternativer Ansatz zur Überprüfung klinischer Kompetenz entwickelt und weltweit übernommen. Ein OSCE besteht aus mehreren Stationen, an denen jeweils spezifische klinische Aufgaben zu erledigen sind. In Deutschland werden seit Änderung der Approbationsordnung für Ärzte im Jahr 2002 zunehmend OSCE-Prüfungen durchgeführt.

Diese Arbeit zielt darauf ab, die aktuelle Situation zu beleuchten. Es soll untersucht werden, inwieweit OSCE-Prüfungen an deutschen medizinischen Fakultäten eingesetzt werden, wie sich der Einsatz der Prüfungsmethode auf das Lernen der Studierenden auswirkt und welche Wahrnehmung die Studierenden hinsichtlich des Nutzens von OSCEs haben.

Mithilfe verschiedener Methoden der Umfrageforschung (Interviews, Fragebogen) wurden Daten im gesamten Bundesgebiet erhoben.

Die Ergebnisse zeigen, dass mittlerweile praktisch alle Fakultäten OSCE-Prüfungen verwenden. Allerdings variieren sowohl das Ausmaß des Einsatzes als auch die Inhalte und Ziele der abgehaltenen OSCEs erheblich zwischen den Fakultäten. Die Ergebnisse zeigen auch, dass die OSCEs einen positiven Effekt auf das Lernverhalten der Studierenden haben. Die Studierenden werden motiviert, sich die notwendigen klinischen Fähigkeiten und Fertigkeiten anzueignen. Die Arbeit zeigt außerdem auf, dass das OSCE-Format von den Studierenden als nützlich für das weitere Lernen und für die Entwicklung klinischer Kompetenz empfunden wird.

Obgleich die OSCE-Methode an den meisten Fakultäten implementiert ist, wird sie vielerorts nur in wenigen Fächern genutzt. Es sollten weitere Anstrengungen unternommen werden, um das Prüfungsformat in der medizinischen Ausbildung der Bundesrepublik Deutschland stärker zu etablieren.

„Je höher der Ausbildungsstandard in einem Beruf ist,
desto geringer ausgeprägt wird die Scharlatanerie sein.“

Sir William Osler

2 Einleitung

Der Berufsstand der Ärzte [schließt sämtliche Geschlechter ein] sieht sich zunehmend hohen gesellschaftlichen Erwartungen gegenüber. Im 21. Jahrhundert werden professionell agierende Ärzte verlangt, die ihr Handeln voll und ganz an den Bedürfnissen der Patienten orientieren. Ein Arzt soll nicht nur fachkundig sein, sondern auch die notwendigen Fähigkeiten, Fertigkeiten und Haltungen im Umgang mit Patienten haben. Um die Studierenden bestmöglich auf ihre zukünftigen Aufgaben vorzubereiten, hat die Gestaltung der medizinischen Ausbildung diesen vielschichtigen Anforderungen Rechnung zu tragen (Gross und Pelz 2009, Medizinischer Fakultätentag 2015).

2.1 Prüfungen in der medizinischen Ausbildung

Prüfungen sind ein elementarer Baustein der medizinischen Ausbildung (Cook et al. 2015) und erfüllen zahlreiche Funktionen: Sie dienen einerseits der Kontrolle des studentischen Lernerfolgs (*Kapitel 2.1.1*) und tragen damit entscheidend zur Qualitätssicherung im Gesundheitswesen bei. Andererseits sind Prüfungen hilfreich, um den Lern- bzw. Entwicklungsprozess der Studierenden zu unterstützen (*Kapitel 2.1.1*). Darüber hinaus haben Prüfungen eine lernsteuernde Wirkung (*Kapitel 2.1.2*) und geben den Fakultäten Rückmeldung über die Effektivität der Lehre (Epstein 2007, Kollwe et al. 2018).

2.1.1 Ziele einer Prüfung

Mit der Durchführung einer Prüfung wird immer ein bestimmtes Ziel verfolgt (Schuwirth und van der Vleuten 2010). Anhand der Zielsetzung lassen sich grundsätzlich zwei Formen von Prüfungen unterscheiden: summative Prüfungen, die in der englischsprachigen Literatur oftmals als „assessment of learning“ bezeichnet werden, und formative Prüfungen bzw. „assessment for learning“ (Wood 2013). Bei der erstgenannten Prüfungsform handelt es sich um eine benotete Erfolgskontrolle, die das Erreichen vorgegebener Lernziele am Ende eines Kurses oder Studiums überprüft. Durch ihre „harte“ abschließende

Bewertung können summative Prüfungen jedoch Entscheidungen beeinflussen, die weitreichende berufliche sowie persönliche Folgen für die Studierenden haben (Sadler 1989). Dagegen stellen formative Prüfungen „weiche“ Instrumente zur Unterstützung des studentischen Lernens dar. Diese Form der Prüfung ist untrennbar mit dem Ausbildungsprozess verknüpft. Im Mittelpunkt steht ein Feedback, um den Studierenden Leistungsdefizite aufzuzeigen und ihnen Hinweise zu geben, wie sie sich verbessern können (Taras 2005, Schuwirth und van der Vleuten 2011, Wood 2013).

2.1.2 Prüfungen und Lernen

Es besteht allgemein Übereinstimmung, dass Prüfungen einen Einfluss auf das Lernen der Studierenden haben (Rust 2002, Cilliers et al. 2010). Sie beeinflussen das Lernen sowohl vor, während als auch nach der Prüfungsaktivität, was als „pre-“, „pure“ und „post-learning effects of assessment“ bezeichnet wird. Die „pre-assessment learning effects“ beziehen sich auf die Prüfungsvorbereitung, wohingegen die „pure“ und „post-assessment effects“ aus der Prüfungserfahrung resultieren bzw. im Zusammenhang mit Feedback und Prüfungsergebnissen stehen (Cilliers et al. 2012, Heeneman et al. 2015). In der vorliegenden Abhandlung geht es hauptsächlich um die Auswirkungen auf das Lernen im Rahmen der Prüfungsvorbereitung („pre-assessment effects“): Der Sinnspruch „assessment drives learning“ ist wohlbekannt (Wood 2009, Newble 2016). Die Studierenden orientieren sich beim Lernen an dem, was geprüft wird. Inhalte, Aufgabenstellungen und Ablauf der Prüfung spielen dabei eine wesentliche Rolle (Cilliers et al. 2012, Kollewe et al. 2018).

2.2 Überprüfung klinischer Kompetenz

Der Begriff „Kompetenz“ wird in der Literatur als ein aus mehreren Komponenten bestehendes Konstrukt zur angemessenen oder erfolgreichen Ausführung einer Tätigkeit beschrieben (Fernandez et al. 2012). Zu der von einem Arzt geforderten professionellen bzw. klinischen Kompetenz zählen zahlreiche Aspekte, wie ein fundiertes medizinisches Wissen (einschließlich dessen Anwendung im klinischen Kontext), praktische Fertigkeiten, kommunikative und zwischenmenschliche Fähigkeiten sowie persönliche Werte (Epstein und Hundert 2002). Angesichts der Komplexität klinischer Kompetenz erfordert

eine sinnvolle Messung den Einsatz valider (aussagekräftiger) und reliabler (zuverlässiger) Prüfungsmethoden (Lee und Wimmers 2011).

Ein konzeptionelles Modell von Miller in Form einer vierstufigen Pyramide umreißt die Thematik (**Abb. 1**). An der Basis der „Miller-Pyramide“ steht Wissen („knows“). Auf der nächsten Stufe folgt die kontextbezogene Anwendung des Wissens („knows how“). Diese beiden Stufen der Kompetenzhierarchie lassen sich durch schriftliche Prüfungen z. B. im Multiple-Choice-(MC-)-Frageformat testen. Gleichwohl erfordert die Praxis der Medizin mehr als den Abruf von Wissen. Miller definiert eine dritte Stufe, wo die Studierenden unter Simulationsbedingungen zeigen sollen, was sie in Konfrontation mit einem Patienten oder einer klinischen Situation tun („shows how“). An die Spitze der Pyramide platziert Miller schließlich die Komponente des Handelns („does“). Diese letzte Stufe der Kompetenzhierarchie zielt darauf abzubilden, wie Prüflinge (Studierende oder Absolventen) im klinischen Alltag agieren (Miller 1990).

Abb. 1: Pyramide nach Miller (1990), die Abbildung wurde aus urheberrechtlichen Gründen entfernt.

2.3 Objective Structured Clinical Examination (OSCE)

Die Methode Objective Structured Clinical Examination oder kurz OSCE ist ein passendes Prüfungsformat für die dritte Stufe der „Miller-Pyramide“ (Davis 2003). Während eines OSCE absolvieren die Studierenden nach einem vorgegebenen Schema eine Reihe von Prüfungsstationen. Jede Station beinhaltet spezifische klinische Aufgabenstellungen und ist mit einem Prüfer besetzt. Die Studierenden verbringen eine definierte Zeit (meistens fünf bis zehn Minuten) an jeder Station und wandern auf ein akustisches Signal hin

zur nächsten Station. Die einzelnen Stationen fokussieren sich auf verschiedene klinische Fähigkeiten bzw. Fertigkeiten, wie Anamneseerhebung, Patientenaufklärung, Untersuchung eines Patienten, Durchführung einer praktischen Maßnahme, Befundauswertung und -beurteilung, Nachrichtenübermittlung, Entscheidungsfindung oder Problemlösung. Um wirklichkeitsnahe Bedingungen bei der Prüfung zu schaffen, kommen geschulte Schauspielpatienten, speziell gefertigte Puppen/Modelle oder echte Patienten zum Einsatz. Prüfer an den jeweiligen Stationen können Ärzte, Studierende oder Schauspielpatienten sein (Harden et al. 2016).

Klassische klinische Prüfungen haben sich als wenig valide und reliabel erwiesen. Kritisiert wurden vor allem die geringe Bandbreite geforderter Kompetenzen sowie die Subjektivität der Bewertung durch die Prüfer (Petruša 2002). Das OSCE-Verfahren, erstmals beschrieben von Harden et al. (1975), wurde entwickelt, um die Probleme der klassischen klinischen Prüfung zu vermeiden.

Der OSCE ist eine performancebasierte Prüfung. Die Studierenden müssen zeigen, dass sie über die notwendigen Kompetenzen für die klinische Praxis verfügen. Das Format kann für eine Vielzahl unterschiedlicher Lernziele, ungeachtet einer bestimmten Disziplin oder Fachrichtung, in jeder Phase der Ausbildung eingesetzt werden. Vor der Durchführung eines OSCE erfolgt auf Basis des studentischen Leistungsstands und der gewünschten Lernziele eine sorgfältige Planung und Festlegung der Prüfungsinhalte. Die Objektivität der Prüfung wird durch eine Verteilung der zu prüfenden Inhalte auf mehrere Stationen und Prüfer sowie durch klar festgelegte Beurteilungskriterien (vorgefertigte Checklisten, Ratingskalen) erreicht. Die Gabe von Feedback an die Studierenden stellt ein wesentliches Element des OSCE dar, vor allem wenn die Prüfung formativ ist (Harden et al. 2016).

Der OSCE hat sich als nützliches Format für summative und formative Prüfungsziele erwiesen und findet seinen Einsatz in der Aus- und Weiterbildung von Medizinern (Townsend et al. 2001, Hodges 2003, Harden 2006, McNaughton et al. 2008). Es konnte gezeigt werden, dass das OSCE-Format eine hohe Validität besitzt (Downing 2003, Roberts et al. 2006). Darüber hinaus wird es auch als reliables Prüfungsverfahren beschrieben (Boursicot et al. 2010). Sowohl bei den Lehrenden als auch den Studierenden stößt die Prüfungsmethode auf breite Akzeptanz (Smith et al. 1984, Ainsworth et al. 1995, Boursicot et al. 2010). In mehreren Ländern ist ein OSCE bereits Bestandteil der

staatlichen Abschluss- bzw. Zulassungsprüfungen (Reznick et al. 1996, Boulet et al. 2009, Kim 2010).

2.4 Ausbildung und Prüfungspraxis in Deutschland

Die medizinische Ausbildung in Deutschland umfasst sechs Jahre bzw. zwölf Semester und unterteilt sich klassischerweise in einen vorklinischen Studienabschnitt (zwei Jahre) und einen klinischen Studienabschnitt (vier Jahre). Die Rahmenbedingungen der Ausbildung werden durch das Bundesministerium für Gesundheit vorgegeben und sind in der aktuellen Fassung der Approbationsordnung für Ärzte (ÄApprO) vom 27.06.2002 geregelt (Approbationsordnung für Ärzte 2002). Die einzelnen Fakultäten haben jedoch weitreichende Spielräume bei der Ausgestaltung ihrer Curricula (Nikendei et al. 2009).

Während der ersten vier Semester (vorklinischer Studienabschnitt) durchlaufen die Studierenden ein naturwissenschaftliches und theoretisches Grundlagenprogramm, das mit dem ersten Staatsexamen (Physikum) abschließt. Danach beginnt der klinische Abschnitt des Studiums. In den ersten drei Jahren dieses Abschnitts (Semester 5 bis 10) werden den Studierenden wesentliche Inhalte klinischer Medizin vermittelt. Für insgesamt 41 Fächer, Querschnittsbereiche und Blockpraktika sind von den Fakultäten benotete Leistungsnachweise zu erstellen (**Tab. 1**). Als Prüfungsinstrumente verwenden die Fakultäten, neben bspw. mündlichen Prüfungen oder Essays, mit Abstand am häufigsten MC-Klausuren. Der Forderung nach einer praxisorientierten Ausbildung entsprechend kommen bei diesen Prüfungen auch zunehmend OSCEs zum Einsatz (Nikendei et al. 2009, Zavlin et al. 2017). Sobald die Studierenden die geforderten Leistungsnachweise für den klinischen Studienabschnitt erworben haben, absolvieren sie normalerweise nach dem zehnten Semester das zweite Staatsexamen. Das beinhaltet eine dreitägige schriftliche Prüfung aus 320 MC-Fragen zur Überprüfung des klinischen Wissens. Hierbei handelt es sich um eine standardisierte Staatsprüfung. Die zu bearbeitenden Aufgaben (Fallstudien und Einzelaufgaben) werden bundeseinheitlich durch das Institut für medizinische und pharmazeutische Prüfungsfragen (IMPP) erstellt. Im Anschluss folgt das Praktische Jahr, das sich aus drei Ausbildungsabschnitten von jeweils 16 Wochen in Innerer Medizin, Chirurgie und einem Wahlfach außerhalb der Inneren Medizin bzw. Chirurgie zusammensetzt. Diese Ausbildungsphase bildet den Übergang zur ärztlichen Berufstätigkeit (Nikendei et al. 2012). Danach findet das dritte Staatsexamen statt. Dieses

Tab. 1: Leistungsnachweise für den klinischen Studienabschnitt gemäß ÄApprO

<i>Fächer</i>	
1.	Allgemeinmedizin
2.	Anästhesiologie
3.	Arbeitsmedizin, Sozialmedizin
4.	Augenheilkunde
5.	Chirurgie
6.	Dermatologie, Venerologie
7.	Frauenheilkunde, Geburtshilfe
8.	Hals-Nasen-Ohrenheilkunde
9.	Humangenetik
10.	Hygiene, Mikrobiologie, Virologie
11.	Innere Medizin
12.	Kinderheilkunde
13.	Klinische Chemie, Laboratoriumsdiagnostik
14.	Neurologie
15.	Orthopädie
16.	Pathologie
17.	Pharmakologie, Toxikologie
18.	Psychiatrie und Psychotherapie
19.	Psychosomatische Medizin und Psychotherapie
20.	Rechtsmedizin
21.	Urologie
22.	Wahlfach
<i>Querschnittsbereiche</i>	
1.	Epidemiologie, medizinische Biometrie und medizinische Informatik
2.	Geschichte, Theorie, Ethik der Medizin
3.	Gesundheitsökonomie, Gesundheitssystem, Öffentliches Gesundheitswesen
4.	Infektiologie, Immunologie
5.	Klinisch-pathologische Konferenz
6.	Klinische Umweltmedizin
7.	Medizin des Alterns und des alten Menschen
8.	Notfallmedizin
9.	Klinische Pharmakologie/Pharmakotherapie
10.	Prävention, Gesundheitsförderung
11.	Bildgebende Verfahren, Strahlenbehandlung, Strahlenschutz
12.	Rehabilitation, Physikalische Medizin, Naturheilverfahren
13.	Palliativmedizin
14.	Schmerzmedizin
<i>Blockpraktika</i>	
1.	Innere Medizin
2.	Chirurgie
3.	Kinderheilkunde
4.	Frauenheilkunde
5.	Allgemeinmedizin

wird an den einzelnen Fakultäten durchgeführt und umfasst eine klinisch-praktische Prüfung über zwei Tage. Am ersten Prüfungstag haben die Studierenden in Gruppen von bis zu vier Kandidaten jeweils einen echten Patienten zu explorieren und daraufhin einen Fallbericht anzufertigen. Der Patient ist nachfolgend einer vierköpfigen Prüfungskommission vorzustellen, die je aus einem Vertreter der Fachgebiete Innere Medizin und Chirurgie, dem im Praktischen Jahr abgeleisteten Wahlfach sowie einer weiteren Disziplin besteht. Dabei werden am Patientenbett klinisch-praktische Fertigkeiten wie Untersuchungs- oder Anamnesetechniken und patientenbezogenes Wissen geprüft. Am zweiten Prüfungstag werden die Studierenden außerhalb des Krankenzimmers durch dieselbe Prüfungskommission erneut mündlich geprüft (Huber-Lang et al. 2017). Ein Bezug zu dem Fallbericht des Vortags wird meistens nicht mehr hergestellt. Die mangelnde Standardisierung der klinisch-praktischen Prüfung (Prüferabhängigkeit, unterschiedliche Patienten) legt nahe, dass dieser Teil der staatlichen Prüfung weniger objektiv und reliabel ist als der schriftliche Anteil. Standardisierte, performancebasierte Formate wie der OSCE sind in den deutschen Staatsprüfungen bisher nicht implementiert.

Derzeit gibt es in Deutschland 38 Fakultäten, an denen nach deutschem Recht Medizin studiert werden kann (Via medici 2017). Eine weitere Fakultät wird gerade in Augsburg aufgebaut (Härtl et al. 2017). Die meisten Fakultäten verfolgen ein traditionelles Curriculum (Regelstudiengang). An einigen Standorten, z. B. an den Fakultäten Berlin, Hannover, Köln oder Mannheim, wurden reformierte Curricula, sog. Modellstudiengänge eingeführt (Nikendei et al. 2009, Via medici 2017). Diese unterscheiden sich für gewöhnlich von den Regelstudiengängen durch eine Aufhebung der strikten Trennung zwischen vorklinischem und klinischem Studienabschnitt. Typisches Merkmal ist, dass das erste Staatsexamen nicht abgelegt werden muss, sondern durch eine gleichwertige Prüfung im Modellstudiengang ersetzt wird (Haage 2002).

3 Ziele der Arbeit

Die vorliegende Arbeit hat zum Ziel, die Verwendung des OSCE in der medizinischen Ausbildung der Bundesrepublik Deutschland zu beleuchten und gegebenenfalls Entwicklungspotenzial aufzuzeigen.

Die Arbeit setzt sich aus drei komplementären Studien zusammen:

Studie I präsentiert eine aktuelle Übersicht zur Implementierung des OSCE an den medizinischen Fakultäten in Deutschland. Vorgestellt wird, wie viele Fakultäten die Prüfungsmethode einsetzen, mit welchem Prüfungsziel (summativ vs. formativ) die OSCEs durchgeführt werden und welche Fächer an der Ausrichtung beteiligt sind bzw. in welche Leistungsnachweise nach ÄApprO die Prüfungsergebnisse einfließen. Dazu wurden alle 36 vor 2012 gegründeten Fakultäten mittels semistrukturierter Telefoninterviews und E-Mail-Korrespondenz befragt.

Studie II veranschaulicht, wie die Einführung von OSCEs das Lernen der Studierenden beeinflusst hat. Mit einem selbst entwickelten Fragebogen wurde untersucht, welche Lernquellen für die Studierenden bei der Vorbereitung auf OSCEs von Bedeutung sind, wie viel Zeit sie zur Vorbereitung aufwenden und welche Prüfungsergebnisse sie erzielen; jeweils im Vergleich zu üblicherweise verwendeten MC-Prüfungen.

Studie III berichtet über den studentisch wahrgenommenen Nutzen des OSCE und der MC-Fragen. Die Studie wurde im Rahmen von Studie II durchgeführt und soll weitere Argumente für eine Einbeziehung des OSCE in die deutschen Staatsprüfungen liefern.

4 Publierte Originalarbeiten

- 4.1 *Objective Structured Clinical Examination (OSCE) an Medizinischen Fakultäten in Deutschland – eine Bestandsaufnahme, S. Müller, U. Dahmen, U. Settmacher, Gesundheitswesen, 80(12):1099–1103, 2018***

Das Gesundheitswesen

DOI: 10.1055/s-0042-116435

Objective Structured Clinical Examination (OSCE) an Medizinischen Fakultäten in Deutschland – eine Bestandsaufnahme

Application of the Objective Structured Clinical Examination (OSCE) in German Medical Schools: An Inventory

S. Müller¹, U. Dahmen², U. Settmacher¹

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

²*Klinik für Allgemein-, Viszeral- und Gefäßchirurgie, Experimentelle Transplantationschirurgie, Universitätsklinikum Jena, Jena*

Schlüsselwörter

Medizinstudium, OSCE-Prüfungsmethode, Fakultäten, Bestandsaufnahme

Keywords

medical education, OSCE assessment, schools, survey

Zusammenfassung

Einleitung: Der Gesetzgeber hat im Jahre 2002 mit Novellierung der Approbationsordnung für Ärzte (ÄApprO) die systematische Bewertung von „Fähigkeiten und Fertigkeiten“ im Medizinstudium verlangt. Daraus resultierte eine sukzessive Einführung von Objective Structured Clinical Examination (OSCE) an den Medizinischen Fakultäten. Die vorliegende Arbeit liefert eine detaillierte Bestandsaufnahme zum derzeitigen Einsatz der OSCE-Prüfungsmethode im Medizinstudium der Bundesrepublik Deutschland. Durch die Erkenntnisse kann der allgemeine sowie fachspezifische Implementierungsbedarf der Prüfungsmethode evaluiert werden.

Methodik: Im Zeitraum von Juni bis September 2015 wurden alle 36 deutschen Medizinischen Fakultäten befragt. Die semistrukturierten Telefoninterviews und Emailkorrespondenzen beinhalteten folgende Themenbereiche zu OSCE: Implementierung, die Anzahl und Art der zu

erwerbenden Leistungsnachweise nach ÄApprO bzw. die Angabe der beteiligten Fachgebiete. Die erhobenen Daten wurden quantitativ und qualitativ analysiert.

Ergebnisse: Alle Medizinischen Fakultäten (100 %) nahmen an der Befragung teil. Fast jede Fakultät (94 %) hat OSCE in ihrem Prüfungsportfolio, allerdings in unterschiedlichem Ausmaß. Bundesweit bilden die Ergebnisse aus OSCE-Prüfungen pro Standort zwischen 0 und 18 ($M = 4,7$) Leistungsnachweise des klinischen Studienabschnitts für die Zulassung zum zweiten Staatsexamen gemäß ÄApprO. Die Implementierung umfasst praktisch alle klinischen Fachgebiete. Dennoch fließt über die Hälfte (51,4 %) der OSCE-Prüfungsergebnisse in Leistungsnachweise für die Fächer Chirurgie, Innere Medizin, Notfallmedizin, Anästhesiologie und Orthopädie.

Schlussfolgerung: Die Ergebnisse dieser Studie bestätigen einen fast flächendeckenden und zugleich in der klinischen Breite verankerten Einsatz der OSCE-Prüfungsmethode. Im Vergleich zwischen den Fakultäten sowie in Bezug auf den Umfang, in welchem einzelne Fachgebiete beteiligt sind, ergibt sich hingegen ein heterogenes Bild. Zur Sicherstellung und dem Nachweis einer umfangreichen klinischen Kompetenz zukünftiger Ärzte sollte der Einsatz von OSCE ausgebaut werden. Hierfür bedarf es noch weiterer Information und Überzeugungsarbeit bei den Lehrverantwortlichen der Fakultäten.

Abstract

Introduction: The German medical licensing regulations, as amended by the government in 2002, now require the assessment of clinical skills in undergraduate medical education. As a result, the Objective Structured Clinical Examination (OSCE) was introduced as an assessment tool in many medical schools. This article provides a detailed inventory of the current application of the OSCE assessment in Germany.

Methods: From June to September 2015, the implementation of the OSCE in all 36 German medical schools was investigated using semi-structured telephone interviews and email correspondence. The areas of focus comprised implementation of the OSCE, and number and type of performance records according to the medical licensing regulations or involved disciplines. Following collection, data were analysed quantitatively and qualitatively.

Results: All medical schools (100 %) participated in the survey. Nearly every school (94 %) has introduced the OSCE into its assessment portfolio, however, to varying extents. For each location, the numbers range between 0 and 18 ($M = 4.7$) performance records of the clinical science section assessed by OSCEs. The implementation of this assessment format includes most of the clinical performance records, but more than half (51.4 %) of these cover surgery, internal medicine, emergency medicine, anaesthesiology, and orthopaedics.

Conclusion: The results reported in this paper confirm the widespread introduction of the OSCE assessment in German medical schools. However, the implementation remains heterogeneous with respect to the scope, schools and individual disciplines involved in the process. In order to ensure extensive clinical competence of prospective physicians the application of the OSCE should be broadened. For this purpose, further information to convince medical school staff is still required.

Einleitung

Zur Sicherung der medizinischen Versorgung werden heutzutage kompetenzorientierte Absolventenprofile als entscheidend angesehen. Für die Klinik gilt die Beherrschung medizinischen Wissens und praktischer Fertigkeiten, verbunden mit professionellen Haltungen, als grundlegende Voraussetzung, um eine optimale Patientenversorgung erreichen zu können [1, 2].

Mit Novellierung der Approbationsordnung für Ärzte (ÄApprO) im Jahre 2002 hat der Gesetzgeber die Notwendigkeit einer systematischen Bewertung von „Fähigkeiten und Fertigkeiten“ im Rahmen des Medizinstudiums betont [3]. Die Vorgaben hatten direkte Auswirkungen auf das Prüfungswesen der Fakultäten. Ein Einsatz innovativer und primär an der Praxis orientierter Prüfungswerkzeuge wurde erforderlich. Daraufhin rückte an den deutschen Medizinischen Fakultäten die in Schottland entwickelte Prüfungsmethode Objective Structured Clinical Examination (OSCE) in den Fokus des Interesses [4, 5].

Im Gegensatz zu traditionellen Prüfungsformaten testet OSCE nicht nur theoretisches Wissen, sondern beurteilt die Studierenden bei der Ausführung ärztlicher Tätigkeiten in klinisch nachempfundenen Situationen. Die Studierenden durchlaufen eine Reihe verschiedener standardisierter Stationen, an denen zahlreiche Komponenten klinischer Kompetenz geprüft werden. Die Inhalte der Prüfungsstationen umfassen angewandtes Wissen und praktische Fertigkeiten (z. B. Anamneseerhebung, körperliche Untersuchung, Stellen von Diagnose und Differenzialdiagnosen, Durchführung kleinerer Interventionen, Problemlösungsstrategien oder Patientenmanagement) in Verbindung mit entsprechenden kommunikativen Fähigkeiten im Umgang mit Patienten (z. B. Authentizität, positive Wertschätzung oder Empathie). Als Patienten kommen trainierte Schauspieler oder standardisierte Patienten zum Einsatz. Für die Beurteilung praktischer Fertigkeiten, die über das Normalmaß einer klinischen Untersuchung hinausgehen, können Plastikmodelle verwendet werden. Die Bewertung der studentischen Prüfungsleistung erfolgt anhand definierter Checklisten [6, 7].

Eine Überprüfung der oben genannten Komponenten klinischer Kompetenz ist im Medizinstudium in der Vergangenheit stark vernachlässigt worden [8, 9]. Die OSCE-Prüfungsmethode hat zum Ziel, diese Lücke zu schließen, indem sie die Performanz der

Studierenden beurteilt [7]. Durch eine Schwerpunktlegung auf psychosoziale Kompetenzen kann OSCE überdies zu einer Verbesserung der in Verruf geratenen Arzt-Patienten-Beziehung beitragen [10].

Zugleich können OSCEs mit 2 unterschiedlichen Intentionen durchgeführt werden. Sie haben entweder summative oder formative Funktion [11]. Summative Prüfungen dienen der Sicherstellung des Vorhandenseins notwendiger Qualifikationen zum Erwerb definierter Leistungsnachweise oder Abschlüsse. Die Prüfungsleistung der Studierenden wird benotet und bildet die Entscheidungsgrundlage für den Fortgang des Studiums. Diese Form der Prüfung hat sanktionierenden Charakter und eignet sich zur Auslese. Formative Prüfungen hingegen liefern eine Standortbestimmung, indem sie Einblicke in den Lernfortschritt der Studierenden geben. Hierdurch lassen sich individuelle Defizite erkennen und konkrete Ausbildungsmaßnahmen können abgeleitet werden. Negative Auswirkungen auf die Fortführung des Studiums haben formative Prüfungen nicht. Die Studierenden bekommen ausschließlich Hinweise über ihren aktuellen Leistungsstand, meistens in Form eines konstruktiven Feedbacks.

Allerdings fehlt eine umfassende Übersicht zum derzeitigen Einsatz der OSCE-Prüfungsmethode in Deutschland. Insbesondere zu dem Ausmaß, in welchem einzelne Fachgebiete zu der Umsetzung beitragen, gibt es bisher kaum Erkenntnisse [12]. Ziel der vorliegenden Studie ist es, eine detaillierte Bestandsaufnahme durchzuführen. Hierdurch kann zur Aufrechterhaltung einer hochwertigen medizinischen Versorgung in der Bundesrepublik Deutschland der allgemeine sowie fachspezifische Implementierungsbedarf der Prüfungsmethode evaluiert werden.

Methodik

Zwischen Juni und September 2015 wurden alle 36 deutschen Medizinischen Fakultäten kontaktiert. Die Datenerhebung erfolgte mittels semistrukturierter Telefoninterviews und Emailkorrespondenz. Als Ansprechpartner fungierten die Studiendekanate der jeweiligen Fakultäten sowie die Lehrbeauftragten der an OSCE-Prüfungen beteiligten Fachgebiete. Der Themenkatalog beinhaltete neben der Erfassung der Form des Studiengangs (Regel- oder Modellstudiengang) verschiedene Aspekte zum fakultätsinternen Einsatz der OSCE-Prüfungsmethode. Die Ansprechpartner wurden gebeten, Aussagen zu Implementierung,

Ausbildungszeitpunkt und Prüfungscharakter (summativ vs. formativ) der OSCEs sowie zu den partizipierenden Fachgebieten zu machen. Ferner sollten die für Studierende durch summative OSCEs semesterweise zu erwerbenden Leistungsnachweise gemäß ÄApprO genannt werden. Bei Fakultäten mit mehreren Studiengängen zur Humanmedizin (parallel angebotener Regel- und Modellstudiengang) wurden sämtliche Daten pro Studiengang erhoben.

Die erhobenen Daten wurden quantitativ analysiert und eine Darstellung von absoluten Häufigkeiten, Prozentsätzen und Mittelwerten erfolgte. Zur Testung auf signifikante Unterschiede zwischen Regel- und Modellstudiengängen diente ein ungepaarter *t*-Test. Wenn eine quantitative Analyse gesammelter Informationen nicht möglich war, wurden diese durch eine qualitative Inhaltsanalyse ausgewertet.

Ergebnisse

Alle 36 Medizinischen Fakultäten (100 %) haben vollständig Auskunft zu den untersuchten Aspekten gegeben. Von diesen 36 Fakultäten haben 34 (94 %) zumindest eine summative oder formative OSCE in ihrem Prüfungsportfolio. Davon setzen 33 Fakultäten (97 %) OSCE summativ ein, während zusätzliche formative OSCEs nur an 10 Standorten (29 %) durchgeführt werden (vgl. **Tab. 1**).

Sämtliche OSCE-Prüfungen werden fast ausschließlich im klinischen Abschnitt bzw. bei Modellstudiengängen in den klinisch geprägten Abschnitten des Studiums veranstaltet. Sie finden entweder in einem fachbezogenen Setting statt, bei dem der gesamte Prüfungsparcours nur von einer Fachdisziplin ausgerichtet wird, oder werden in einem multidisziplinären Rahmen abgehalten. Hierbei können bis zu 9 verschiedene Voll- oder Teilleistungsnachweise erworben werden (summative Form) bzw. maximal 14 unterschiedliche Fachgebiete sind gegenwärtig in einem solchen Parcours mit Prüfungsstationen vertreten (formative Ausrichtung). **Tab. 2** stellt eine Auswahl angewandter Fächerkombinationen bei multidisziplinären OSCEs dar.

Bundesweit setzen die Fakultäten summative OSCE-Prüfungen in 185 Fällen zum Erbringen von Leistungsnachweisen für die Zulassung zum zweiten Staatsexamen (klinischer Studienabschnitt) gemäß § 27 ÄApprO ein. Die Einsatzhäufigkeit differiert jedoch erheblich zwischen den Standorten, wobei Lehrforschung betreibende Fakultäten die Prüfungsmethode häufiger einsetzen. Die Spannweite reicht von 0 bis 18 Leistungsnachweisen ($M = 4,7$), die pro

Studiengang aus OSCE-Prüfungsergebnissen gebildet werden. Differenziert nach der Form des Studiengangs lassen sich keine signifikanten Unterschiede in der Häufigkeit nachweisen ($p = 0,37$). In den gesamten Regelstudiengängen ($n = 31$) bilden OSCEs in 136 Fällen ($M = 4,4$) die Grundlage für den Erwerb von Leistungsnachweisen des klinischen Studienabschnitts. Bei den Modellstudiengängen ($n = 8$; bei einer Fakultät war keine eindeutige Zuordnung möglich) tragen sie dazu in insgesamt 49 Fällen ($M = 6,1$) bei. Darüber hinaus haben 7 der 9 Modellstudiengänge (78 %) das OSCE-Format in ihre physikumsäquivalenten Prüfungen (erstes Staatsexamen ersetzend) integriert, was bei den regulären Staatsexamina nicht vorgesehen ist. Im Rahmen des Erwerbs von Leistungsnachweisen für die Zulassung zum ersten Staatsexamen (vorklinischer Studienabschnitt) wird die OSCE-Prüfungsmethode indessen bundesweit in nur 5 Fällen eingesetzt.

Die Durchführung summativer OSCEs betrifft praktisch alle Leistungsnachweise des klinischen Studienabschnitts. Die Häufigkeit des Einsatzes ist jedoch zwischen den verschiedenen Leistungsnachweisen unterschiedlich stark ausgeprägt. Derzeit fließen die Ergebnisse aus OSCEs fast zu einem Drittel (31,4 %) in Nachweise für die beiden großen Fächer Chirurgie (17,3 %) und Innere Medizin (14,1 %). Zugleich findet sich ein häufiger Eingang der Prüfungsergebnisse in Nachweise für die Fachgebiete Notfallmedizin (7,6 %), Anästhesiologie (6,5 %) und Orthopädie (5,9 %). Für das Erbringen von Leistungsnachweisen in einer Vielzahl anderer klinischer Fächer haben summative OSCEs dagegen kaum Bedeutung (vgl. **Tab. 3**). Formative OSCEs sind überwiegend an Standorten mit traditionellem Regelstudiengang implementiert. Obgleich diese Form der OSCE-Prüfung gegenüber der summativen deutlich seltener durchgeführt wird, zeigt sich hinsichtlich der beteiligten Fachgebiete ein ähnliches Bild (vgl. **Tab. 2**).

Diskussion

Die Ergebnisse dieser Arbeit belegen einen mittlerweile fast flächendeckenden Einsatz der OSCE-Prüfungsmethode in Deutschland. Demgegenüber waren im Jahr 2008 nur an 28 (78 %) der bundesdeutschen Fakultäten klinisch-praktische Prüfungen im OSCE-Format implementiert [12]. Im Vergleich zwischen den Fakultäten ergibt sich ein heterogenes Bild. Die Prüfungsmethode kommt vorrangig an Standorten mit intensiver Lehrforschung und der

Implementierung innovativer Curricula zum Einsatz. Die betreffenden Fakultäten haben oftmals eigene Einrichtungen für Ausbildungsforschung oder kooperieren mit entsprechenden Kompetenzzentren. Eine Vielzahl der Lehrverantwortlichen besitzt Zusatzqualifikationen wie Master of Medical Education (MME). Diese Errungenschaften haben offenbar einen positiven Einfluss auf die Qualität und Fortschrittlichkeit des Prüfungswesens. Die Form des Studiengangs scheint hingegen, abgesehen von den das erste Staatsexamen ersetzenden Prüfungen, bei der Implementierung von OSCE keine Rolle zu spielen.

Zwischenzeitlich hat sich die OSCE-Prüfungsmethode auch in der klinischen Breite etabliert. Die Implementation betrifft nicht mehr nur einzelne Fächer [12], sondern umfasst nahezu das gesamte Spektrum klinischer Fachgebiete. In Bezug auf den Umfang des Einsatzes kann jedoch eine erhebliche Spannweite nachgewiesen werden. Neben den beiden großen Fächern Chirurgie und Innere Medizin ragen vor allem die Fachgebiete Notfallmedizin, Anästhesiologie und Orthopädie hervor. Eine Dominanz dieser Fachgebiete konnte bereits von Möltner et al. (2010) gezeigt werden [13]. Aufgrund des starken Praxisbezuges sind OSCEs für solche Fächer, die einen hohen klinisch-praktischen Anteil aufweisen, natürlich besonders zweckmäßig [14, 15]. Die Prüfungsmethode ist jedoch variabel und grundsätzlich für alle Fächer anwendbar [16, 17]. Insofern besteht bei den unterrepräsentierten Fachgebieten offenkundig noch ein gewisser Nachholbedarf. Wohlgermerkt wird diesen meist kleinen Fächern, wie Arbeitsmedizin oder Dermatologie, im Gegensatz zu den 5 Fachgebieten mit durch die ÄApprO vorgeschriebenen Blockpraktika (Innere Medizin, Chirurgie, Kinderheilkunde, Frauenheilkunde, Allgemeinmedizin) weniger Platz in den Curricula eingeräumt. Folglich haben diese weniger die Möglichkeit, ihre Inhalte zu lehren und zu prüfen. Allerdings eignen sich die multidisziplinär angelegten OSCE-Prüfungsformate hervorragend zur Beteiligung für diese Fächer, die auch aufgrund geringerer finanzieller und personeller Ressourcen oftmals keine eigenständige OSCE ausrichten können. Als angenehmem Nebeneffekt wird dabei zugleich der Forderung einer horizontalen und vertikalen Verknüpfung der Prüfungsinhalte gemäß der ÄApprO Rechnung getragen [3].

Formative OSCE-Prüfungen werden gegenüber summativen deutlich seltener veranstaltet. Diese Form der Beurteilung findet angesichts der hohen Prüfungsdichte im Medizinstudium

kaum Anklang bei den Studierenden. Daher planen manche Fakultäten, ihre bestehenden formativen OSCEs in summative Prüfungen umzuwandeln. Bei summativen OSCEs ist ein Feedback für den Fortschritt des Lernerfolges indes nicht regelhaft vorgesehen. Prüfungen sollten jedoch nicht nur summative, sondern auch eine die Studierenden prägende Funktion haben. Sie müssen lehrreich sein [18], denn nur durch die Überprüfung von Wissen und Kompetenzen sowie einer geeigneten Rückmeldung dazu kann das Lernen sinnvoll unterstützt werden [19]. Dementsprechend sollte ein strukturiertes Feedback unbedingt in die summativen OSCEs integriert werden.

Die vorgestellten Zahlen zur Verwendung der OSCE-Prüfungsmethode bestätigen eine positive Entwicklung, die allerdings bei Weitem noch nicht zufriedenstellend ist. Augenblicklich spielt OSCE weiterhin nur eine untergeordnete Rolle in der Prüfungskultur des Medizinstudiums. Fehlendes Wissen und mangelnde Erfahrung der Lehrverantwortlichen mit der Prüfungsmethode mögen Ursachen hierfür sein. Zudem besitzt die Lehre keine große Anerkennung im akademischen Umfeld, was in einer Zurückhaltung seitens der Dozenten gegenüber dieser Prüfungsform resultiert [16]. Allerdings gibt es auch handfeste Gründe, die gegen einen Einsatz der OSCE-Prüfungsmethode sprechen. Ein hoher organisatorischer und planerischer Aufwand sowie die Entstehung beträchtlicher Kosten sind Nachteile der Prüfungsmethode [14, 20]. Dies gilt insbesondere für summative OSCEs, die im Hinblick auf eine rechtssichere Prüfungsorganisation ein Mindestmaß an Qualitätsmerkmalen, wie Objektivität, Reliabilität und Validität (Testgütekriterien), erfüllen müssen [4]. Gleichwohl gibt es gegenwärtig an mehreren Fakultäten Vorhaben, in den nächsten Jahren weitere OSCE-Prüfungen zu installieren.

Zur Gewährleistung und der Bescheinigung einer umfangreichen klinischen Kompetenz angehender Ärzte ist ein forcierter Ausbau des Einsatzes der OSCE-Prüfungsmethode, vornehmlich in den bislang unterrepräsentierten klinischen Fachgebieten, anzustreben. Dafür bedarf es allerdings noch einiger Information und Überzeugungsarbeit bei den Lehrverantwortlichen der Fakultäten, was beispielsweise im Rahmen von Fortbildungen oder Schulungen erfolgen kann. Gleichzeitig können Qualifikationsprogramme wie MME einen wichtigen Beitrag leisten, um die Lehre für Ärzte attraktiver zu machen. Zweifelsohne ist der

Einsatz qualitätssichernder OSCE-Prüfungen, der eine sorgfältige Planung und Durchführung sowie eine Mindestanzahl an Prüfungsstationen erfordert [21], sehr ressourcenintensiv. Angesichts der zahlreichen Vorteile, die die OSCE-Prüfungsmethode bietet, sollten die Fakultäten diese Aufwendungen jedoch in Kauf nehmen und zugleich versuchen, die Durchführung ihrer OSCEs so kosteneffizient wie möglich zu gestalten [20]. Nachfolgende Erhebungen werden zeigen, wie sich die derzeitige Entwicklung der Implementierung in Zukunft fortsetzt.

Literatur

- 1 *Fischer MR*. Medizinstudium als Grundstein für Versorgung und Forschung. Dtsch Arztebl 2012; 109: 325–326
- 2 *Newble DJ*. Assessing clinical competence at the undergraduate level. Med Educ 1992; 26: 504–511
- 3 *Bundesministerium für Gesundheit*. Approbationsordnung für Ärzte. ÄApprO 2002
- 4 *Schultz JH, Nikendei C, Weyrich P et al*. Qualitätssicherung von Prüfungen am Beispiel des OSCE-Prüfungsformats: Erfahrungen der Medizinischen Fakultät der Universität Heidelberg. Z Evid Fortbild Qual Gesundheitswesen 2008; 102: 668–672
- 5 *Harden RM, Stevenson M, Downie WW et al*. Assessment of clinical competence using objective structured examination. Br Med J 1975; 1: 447–451
- 6 *Harden RM, Gleeson FA*. Assessment of clinical competence using an objective structured clinical examination (OSCE). Med Educ 1979; 13: 41–54
- 7 *Harden RM*. What is an OSCE? Med Teach 1988; 10: 19–22
- 8 *Jungbauer J, Kamenik C, Alfermann D et al*. Wie bewerten angehende Ärzte rückblickend ihr Medizinstudium? Ergebnisse einer Absolventenbefragung. Gesundheitswesen 2004; 66: 51–56
- 9 *Hofer M, Jansen M, Soboll S*. Verbesserungspotenzial des Medizinstudiums aus retrospektiver Sicht von Facharztprüflingen. Dtsch Med Wochenschr 2006; 131: 373–378
- 10 *Thielscher C, Schulte-Sutrum B*. Die Entwicklung der Arzt-Patienten-Beziehung in Deutschland in den letzten Jahren aus Sicht von Vertretern der Ärztekammern und der Kassenärztlichen Vereinigungen. Gesundheitswesen 2016; 78: 8–13

- 11 *Nikendei C, Jünger J.* OSCE – praktische Tipps zur Implementierung einer klinisch-praktischen Prüfung. *GMS Z Med Ausbild* 2006; 23: Doc47
- 12 *Kruppa E, Jünger J, Nikendei C.* Einsatz innovativer Lern- und Prüfungsmethoden an den Medizinischen Fakultäten der Bundesrepublik Deutschland. Eine aktuelle Bestandsaufnahme. *Dtsch Med Wochenschr* 2009; 134: 371–372
- 13 *Möltner A, Duelli R, Resch F et al.* Fakultätsinterne Prüfungen an den deutschen medizinischen Fakultäten. *GMS Z Med Ausbild* 2010; 27: Doc44
- 14 *Weißer FO, Dirks B, Georgieff M.* Objective Structured Clinical Examination (OSCE). Eine neue Prüfungsform in der notfallmedizinischen Ausbildung. *Notfall Rettungsmed* 2004; 7: 237–243
- 15 *Kalbitz M, Liener U, Kornmann M et al.* Studentische Evaluation einer objektiven, strukturierten klinischen Prüfungsmethode (OSCE) im Fach Chirurgie und Orthopädie. *Unfallchirurg* 2010; 113: 726–733
- 16 *Brauer RB, Kammerloher A, Stering K et al.* Implementierung einer OSCE-Prüfung zur Benotung des Chirurgischen Blockpraktikums. Zwölf Monate Erfahrung mit einer praktischen Prüfungsmethode. *Zentralbl Chir* 2013; 138: 144–150
- 17 *Simmenroth-Nayda A, Görlich Y, Wagner M et al.* Studentische Lehre in der Augenheilkunde. Sind standardisierte praktische Prüfungen sinnvoll? *Ophthalmologe* 2014; 111: 235–240
- 18 *Wass V, van der Vleuten C, Shatzer J et al.* Assessment of clinical competence. *Lancet* 2001; 357: 945–949
- 19 *Fischer MR, Holzer M, Jünger J.* Prüfungen an den medizinischen Fakultäten. Qualität, Verantwortung und Perspektiven. *GMS Z Med Ausbild* 2010; 27: Doc66
- 20 *Brown C, Ross S, Cleland J et al.* Money makes the (medical assessment) world go round: The cost of components of a summative final year objective structured clinical examination (OSCE). *Med Teach* 2015; 37: 653–659
- 21 *Barman A.* Critiques on the objective structured clinical examination. *Ann Acad Med Singapore* 2005; 34: 478–482

Tab. 1 Übersicht zur Implementierung der OSCE-Prüfungsmethode an den Medizinischen Fakultäten in Deutschland ($n = 36$); Anteile der Standorte mit Implementation des jeweiligen Prüfungscharakters.

OSCE	Anzahl der Fakultäten $n / 36$ (%) mit Implementation
Gesamt	34 (94 %)
nur summativ	24 (71 %)
summativ und formativ	9 (26 %)
nur formativ	1 (3 %)

Tab. 2 Darstellung einer Auswahl angewandter Fächerkombinationen bei multidisziplinären OSCE-Prüfungen (summativ bzw. formativ).

OSCE	Fächerkombination
summativ	<ul style="list-style-type: none"> • Anästhesiologie; Chirurgie; Innere Medizin; Notfallmedizin; Orthopädie; Urologie
	<ul style="list-style-type: none"> • Chirurgie; Frauenheilkunde; Innere Medizin; Kinderheilkunde
	<ul style="list-style-type: none"> • Allgemeinmedizin; Chirurgie; Innere Medizin
	<ul style="list-style-type: none"> • Allgemeinmedizin; Geriatrie; Psychiatrie; Psychosomatische Medizin
	<ul style="list-style-type: none"> • Chirurgie; Dermatologie; Innere Medizin; Kinderheilkunde; Neurologie; Orthopädie; Psychosomatische Medizin; Radiologie
	<ul style="list-style-type: none"> • Allgemeinmedizin; Arbeits- und Sozialmedizin; Chirurgie; Dermatologie; Hygiene/Mikrobiologie; Innere Medizin; Notfallmedizin; Psychosomatische Medizin; Radiologie
formativ	<ul style="list-style-type: none"> • Chirurgie; Hals-Nasen-Ohrenheilkunde; Hygiene/Mikrobiologie; Innere Medizin; Klinische Chemie; Pathologie; Pharmakologie
	<ul style="list-style-type: none"> • Allgemeinmedizin; Augenheilkunde; Anästhesiologie; Chirurgie; Hals-Nasen-Ohrenheilkunde; Hygiene/Mikrobiologie; Innere Medizin; Kinderheilkunde; Medizinische Psychologie/Soziologie; Neurologie; Orthopädie; Palliativmedizin; Psychiatrie; Rechtsmedizin
	<ul style="list-style-type: none"> • Neurologie; Psychiatrie; Psychosomatische Medizin • Allgemeinmedizin; Augenheilkunde; Chirurgie; Dermatologie; Frauenheilkunde; Hals-Nasen-Ohrenheilkunde; Innere Medizin; Kinderheilkunde; Neurologie; Notfallmedizin; Orthopädie; Psychiatrie; Psychosomatische Medizin; Urologie

Tab. 3 Differenzierung summativer OSCEs nach erwerbbaeren Leistungsnachweisen (Fächer, Querschnittsbereiche, Blockpraktika) gemäß § 27 ÄApprO.

Leistungsnachweis	Häufigkeit des Erwerbs <i>n</i> / 185 (%) durch OSCE *
Allgemeinmedizin	3 (1,6 %)
Anästhesiologie	12 (6,5 %)
Arbeitsmedizin, Sozialmedizin	1 (0,5 %)
Augenheilkunde	3 (1,6 %)
Chirurgie	16 (8,6 %)
Dermatologie, Venerologie	2 (1,1 %)
Frauenheilkunde, Geburtshilfe	2 (1,1 %)
Hals-Nasen-Ohrenheilkunde	6 (3,2 %)
Hygiene, Mikrobiologie, Virologie	5 (2,7 %)
Innere Medizin	15 (8,1 %)
Kinderheilkunde	2 (1,1 %)
Klinische Chemie, Laboratoriumsdiagnostik	2 (1,1 %)
Neurologie	5 (2,7 %)
Orthopädie	11 (5,9 %)
Pathologie	4 (2,2 %)
Pharmakologie, Toxikologie	5 (2,7 %)
Psychiatrie und Psychotherapie	2 (1,1 %)
Psychosomatische Medizin und Psychotherapie	7 (3,8 %)
Urologie	6 (3,2 %)
Klinisch-pathologische Konferenz	2 (1,1 %)
Medizin des Alterns und des alten Menschen	1 (0,5 %)
Notfallmedizin	14 (7,6 %)
Klinische Pharmakologie/Pharmakotherapie	1 (0,5 %)
Bildgebende Verfahren, Strahlenbehandlung, Strahlenschutz	2 (1,1 %)
Rehabilitation, Physikalische Medizin, Naturheilverfahren	1 (0,5 %)
Palliativmedizin	2 (1,1 %)
Schmerzmedizin	2 (1,1 %)
BP Innere Medizin	11 (5,9 %)
BP Chirurgie	16 (8,6 %)
BP Kinderheilkunde	8 (4,3 %)
BP Frauenheilkunde	8 (4,3 %)
BP Allgemeinmedizin	6 (3,2 %)
Sonstige **	2 (1,1 %)

* bestandene OSCE-Prüfung teilweise Zugangsvoraussetzung für den Erwerb des entsprechenden Leistungsnachweises; ** fakultätsinterne

Anpassung der von § 27 ÄApprO geforderten Leistungsnachweise; BP, Blockpraktikum

4.2 *How the introduction of OSCEs has affected the time students spend studying: results of a nationwide study, Stefan Müller, Ines Koch, Utz Settmacher and Uta Dahmen, BMC Med Educ, 19:146, 2019*

RESEARCH ARTICLE

Open Access



How the introduction of OSCEs has affected the time students spend studying: results of a nationwide study

Stefan Müller¹, Ines Koch², Utz Settmacher¹ and Uta Dahmen^{3*} 

Abstract

Background: Medical schools globally now use objective structured clinical examinations (OSCEs) for assessing a student's clinical performance. In Germany, almost all of the 36 medical schools have incorporated at least one summative OSCE into their clinical curriculum. This nationwide study aimed to examine whether the introduction of OSCEs shifted studying time. The authors explored what resources were important for studying in preparation for OSCEs, how much time students spent studying, and how they performed; each compared to traditionally used multiple choice question (MCQ) tests.

Methods: The authors constructed a questionnaire comprising two identical sections, one for each assessment method. Either section contained a list of 12 study resources requesting preferences on a 5-point scale, and two open-ended questions about average studying time and average grades achieved. During springtime of 2015, medical schools in Germany were asked to administer the web-based questionnaire to their students in years 3–6. Statistical analysis compared the responses on the open-ended questions between the OSCE and MCQs using a paired *t*-test.

Results: The sample included 1131 students from 32 German medical schools. Physical examination courses were most important in preparation for OSCEs, followed by class notes/logs and the skills lab. Other activities in clinical settings (e.g. medical clerkships) and collaborative strategies ranked next. Conversely, resources for gathering knowledge (e.g. lectures or textbooks) were of minor importance when studying for OSCEs. Reported studying time was lower for OSCEs compared to MCQ tests. The reported average grade, however, was better on OSCEs.

Conclusions: The study findings suggest that the introduction of OSCEs shifted studying time. When preparing for OSCEs students focus on the acquisition of clinical skills and need less studying time to achieve the expected level of competence/performance, as compared to the MCQ tests.

Keywords: Assessment, OSCEs, Study behaviour, Medical students, Survey

Background

Medical schools around the world have implemented objective structured clinical examinations (OSCEs) [1]. In an OSCE, students move through a series of stations where they have to perform specific clinical tasks within a time limit. The content domains to be assessed and the scoring scheme for the examination are defined in advance [2]. Since its first description in the mid-1970s [3], the OSCE has been the subject of countless papers

[4]. A number of papers have shown that the OSCE is a valid and reliable assessment of a student's clinical competence [5–8]. Papers have also shown that students accept the OSCE as a relevant and fair exam [9–11].

However, only a few studies exist on how the deployment of OSCEs affects students' study behaviour. Newble and Jaeger [12], for instance, reported that work-based learning, textbooks, tutorials, and group activities were the predominant resources when studying for a clinical examination. Mavis [13] found that students focused on cognitive learning strategies, such as reviewing textbooks or class notes, when preparing for an OSCE. This study, however, was limited to the extent that the

* Correspondence: uta.dahmen@med.uni-jena.de

³Department of General, Visceral and Vascular Surgery, Experimental Transplantation Surgery, Universitätsklinikum Jena, Drackendorfer Str. 1, 07747 Jena, Germany

Full list of author information is available at the end of the article



examined OSCE was a formative, and not a summative assessment, which may explain the different findings. Rudland and colleagues [14] identified that the OSCE fostered collaborative learning, but did not encourage students to spend more time learning in clinical settings. The disparities found in the literature suggest that the OSCE does not always drive student learning in the desirable way. The student study behaviour may rather depend upon of what is specifically assessed in the OSCEs, the purpose of the assessment (summative vs. formative), as well as other factors such as patient availability [15], advice given by the teachers or information from peers.

The aim of our study was to examine whether the introduction of OSCE assessments shifted the time students spend studying. We explored what resources were important for studying and how much time students spent when studying for OSCEs compared to traditionally used multiple choice question (MCQ) tests, and how they performed on the respective assessment format.

Methods

Context

We conducted the present study in the context of the amendment of the national medical licensure act carried out in 2002, which called for a more practice- and patient-oriented alignment of medical education in Germany [16]. Each of the 36 medical schools established before 2012 has a six-year curriculum. The curricula usually consist of two preclinical years followed by three clinical years and, finally, the clinical internship year. According to the guidelines of the medical licensure act, the clinical years comprise 41 predetermined courses entailing the full range of clinical areas or disciplines. During these courses, students have to pass summative (graded) in-house assessments designed by medical school members to be admitted to the national licensing examination. The medical licensure act sets the general framework of the undergraduate programme, but schools have considerable freedom to organise their own curricula. Thus, both the succession of the individual courses and their specific content, as well as the accompanying assessment strategy differ from one school to another.

Written assessments, in the form of MCQ tests, are still most commonly used during the clinical years in all German medical schools. The focus is on testing a student's knowledge about diseases, involving pathogenesis, signs and symptoms, diagnostic approaches, and treatment strategies. In order to comply with the new legal requirements, medical schools have broadened their assessment repertoire to include performance related skills. By now, 33 of the 36 medical schools (92%)

have introduced at least one summative OSCE into their in-house assessment system used for the clinical curriculum [17]. In the held OSCEs, the main focus is on the performance domains physical examination, history taking, practical procedures, and communication skills. Passing the OSCE(s) is also a prerequisite for students to be admitted to the national licensing examination.

Student population

In the academic year 2014/15, there were around 88,000 medical students in Germany. Almost 2/3 of them (53,352 [61%]) were female students [18]. We surveyed medical schools on both the number of students per year and the timing when OSCEs occurred in the curriculum. With these data, we calculated the proportion of students during the clinical years or the clinical internship year who had exposure to a summative OSCE at slightly more than 32,700.

Data collection

Between February and April of 2015, we conducted this study using the free software package SoSci Survey (www.sosicisurvey.de). Due to privacy terms, we did not get access to the students' email addresses. We therefore could not administer our web-based questionnaire to a selected sample of the population in study; but instead, asked the medical schools in Germany to advertise the survey on their websites or through messaging systems. All medical students of years 3, 4, 5, and 6, who had undertaken at least one summative OSCE, were eligible to participate in the study. Participation was voluntary and anonymous, and the respondents did not receive any incentive for completing the questionnaire. The study was in accordance with the ethical standards of our institutional review board (Ethics Committee of Jena University Hospital at Friedrich Schiller University).

Design of the questionnaire

We first reviewed literature and conducted interviews with students to identify items that we could use for our study. Based on this knowledge, we developed a draft questionnaire. As a next step, we repeatedly pilot-tested and revised the draft for ensuring that respondents completed the survey in the intended manner. The final version of the questionnaire comprised two identical sections, the first for the OSCE and the second for the MCQs.

Each section contained a list of 12 study resources (Table 1). Participants rated their preferences in preparation for the respective assessment method on a 5-point scale, anchoring 1 (*not important*), 2 (*slightly important*), 3 (*moderately important*), 4 (*important*), and 5 (*very important*). Participants then answered two open-ended questions. Firstly, we prompted them to

Table 1 Students' preferences of study resources when preparing for OSCEs and MCQ tests

Study resource	OSCEs	MCQ tests
	Mean (SD)	Mean (SD)
Physical examination courses	4.38 (0.82)	2.09 (1.07)
Class notes/logs	3.88 (1.11)	4.25 (0.95)
Skills lab	3.87 (1.21)	1.83 (0.97)
Group learning	3.61 (1.26)	2.57 (1.29)
Medical clerkships	3.58 (1.16)	2.48 (1.14)
Clinical work placements	3.56 (1.12)	2.37 (1.13)
Peer tutorials	3.53 (1.19)	2.15 (1.11)
Textbooks	3.17 (1.07)	3.97 (0.98)
Multimedia materials	3.05 (1.11)	3.32 (1.19)
Casebooks	2.90 (1.14)	2.42 (1.14)
Lectures	2.67 (1.08)	4.07 (1.03)
PBL courses	2.48 (1.14)	2.12 (1.09)

PBL problem-based learning, SD standard deviation

The responses were made on a 5-point scale with anchors 1 (*not important*), 2 (*slightly important*), 3 (*moderately important*), 4 (*important*), and 5 (*very important*)

Number of respondents, $n = 1131$

estimate the average total time they spent preparing for a single summative in-house OSCE or MCQ test. To improve ease of completion, we requested them to indicate their time spent in working days of about 8 h. Secondly, we asked them to report their overall average grade achieved on each of the two assessment methods. In this paper, we present the reported grades on a 4.0 grading scale ranging from 0.0 (failing grade) and 1.0 (lowest passing grade) to 4.0 (best possible grade). Finally, the questionnaire collected demographic details, involving gender, age, academic year (semester), and medical school affiliation.

The questionnaire also included an 11-item set on the benefit of the OSCE or MCQs (at the beginning of each section). The results are presented in a separate paper recently published in GMS Journal for Medical Education [19]. All the questionnaire items can be found in Additional file 1.

Data analysis

After sampling, we verified that each respondent had exposure to a summative OSCE by squaring the indicated semester or the day when completing our questionnaire with the specific curriculum of the relevant medical school. For carrying out the statistical analysis, we used IBM SPSS Statistics for Windows, Version 24 (IBM Corp., Armonk, NY, USA). We performed descriptive statistics and used a paired t -test to compare participants' responses on the open-ended questions between the OSCE and MCQs. We calculated Cohen's d as a measure of effect size from the t -statistic (t -value, group

size, and Pearson's correlation coefficient). To determine whether the medical school had an influence on the results, we conducted a univariate ANOVA using the mean of the responses for the two assessment methods (excluding missing data) as the dependent variable and the medical school as the fixed factor. We considered p values below 0.05 statistically significant. When not stated otherwise, we present data as means with standard deviations in parentheses.

Results

Sample

The number of participants who completed the questionnaire was 1189. We removed 58 respondents, as either the demographic details were incomplete or we observed that those respondents had not yet been exposed to a summative in-house OSCE. Our analysis included 1131 respondents (777 female students [69%] vs. 354 male students [31%]) from 32 of the 33 medical schools (97%) that were holding summative OSCE(s). The sample represented all age groups of students, with a range from 19 to 45 years (median 25). Group sizes of years 4, 5, and 6 were similar (318 [28%], 338 [30%], and 303 [27%], respectively), while the proportion of year 3 due to less exposure to OSCEs was lower (172 [15%]). The number of respondents in each of the 33 medical schools varied between 0 and 123, with a mean of 34 responses per school, depending on how the individual schools advertised our study (websites vs. messaging systems).

Study resources

All 1131 respondents included into the analysis completed the list on study resources. The ratings indicated that in preparation for OSCEs, students mostly preferred resources to acquire clinical skills. Physical examination courses (4.38 [0.82]) ranked first, followed by class notes/logs (3.88 [1.11]) and the skills lab (3.87 [1.21]). Medical clerkships (3.58 [1.16]) and clinical work placements (3.56 [1.12]), as well as group learning (3.61 [1.26]) and peer tutorials (3.53 [1.19]) ranked next. The ratings also showed that students attached moderate importance to resources for gathering knowledge, such as lectures (2.67 [1.08]) or textbooks (3.17 [1.07]), when studying for OSCEs.

Students' preferences of resources for studying were different when they were preparing for MCQ tests. Resources to gather knowledge were most important, whereas those to acquire clinical skills were of minor importance. Class notes/logs (4.25 [0.95]), lectures (4.07 [1.03]), and textbooks (3.97 [0.98]) ranked first, second, and third, followed by multimedia materials (3.32 [1.19]). Table 1 shows the complete ratings on the list of study resources for both assessment methods.

Studying time

We obtained valid responses from 1043 respondents. The reported number of hours spent for an OSCE was 66.5 (52.5) and 94.8 (71.5) for an MCQ test, respectively, which was significantly different ($t[1042] = -14.78, p < 0.01$). Cohen's d of 0.44 showed an effect size in the medium range.

The ANOVA revealed that the medical school had a significant influence on the duration of studying ($F[31, 1011] = 5.40, p < 0.01$, partial eta squared = 0.14). Table 2 includes the results on studying time by medical school affiliation. We found that respondents from about half of the medical schools (15/32 [47%]) reported a significantly lower time spent in preparation for an OSCE compared to an MCQ test, while the time spent did not differ significantly between the assessment methods for respondents of the other schools.

Performance outcomes

From our respondents, 1111 replied to the question about performance outcomes. The reported average grade was 3.13 (0.62) on OSCEs and 2.84 (0.62) on MCQ tests, respectively. The difference was significant ($t[1110] = 12.55, p < 0.01$). Cohen's d effect size was 0.47 indicating a medium effect.

There was a significant influence of the medical school on the OSCE and MCQ grades ($F[31, 1079] = 6.48, p < 0.01$, partial eta squared = 0.16). Table 3 shows the results on performance outcomes by medical school affiliation. Our analysis revealed that the reported grades were significantly better on OSCEs for respondents of almost half of the medical schools (14/32 [44%]), whereas the grades on MCQ tests were significantly better only in one school (XV).

Discussion

In response to the amendment of the national licensure act, German medical schools have incorporated OSCEs into their system of assessment. This nationwide study sought to address how the introduction of OSCEs has affected the time students spend studying. We identified that students use different strategies to prepare in advance of OSCE assessments than common MCQ tests. However, this finding was not surprising: When preparing for an assessment, students adapt their study behaviour (what and how they learn) to the assessment rather than to the learning objectives laid down in the curriculum. Both the content domains to be expected and the tasks required in the upcoming assessment influence student learning [20, 21], which has been described as pre-assessment learning effects of assessment [22]. Given the tasks being tested in the OSCEs (taking a history, examining a patient or carrying out a procedure), we therefore expected that students seek opportunities to rehearse the desired clinical skills.

Table 2 Reported studying time for OSCEs and MCQ tests by medical school affiliation

Medical school	n^b	Time spent (h) for an OSCE ^a	Time spent (h) for an MCQ test ^a	p^c
		Mean	Mean	
VI	51	51.8	87.7	< 0.01
VII	38	41.6	76.5	< 0.01
VIII	38	50.6	109.5	< 0.01
X	108	89.1	146.2	< 0.01
XI	34	51.7	92.2	< 0.01
XII	32	54.3	120.6	< 0.01
XIV	72	59.9	82.0	< 0.01
XVI	40	52.5	105.8	< 0.01
XXXI	55	46.4	80.0	< 0.01
II	11	54.5	103.3	0.01
III	17	57.2	84.7	0.01
XXII	8	85.0	156.0	0.02
XXIII	18	44.1	95.6	0.02
XXIV	38	111.8	139.8	0.02
XXXII	50	47.9	63.7	0.04
IX	33	67.9	89.7	0.06
XXX	25	54.2	70.4	0.06
V	16	45.5	59.5	0.07
I	18	53.8	79.8	0.08
XXV	6	60.0	89.3	0.08
XXI	51	56.9	69.6	0.09
XXVIII	54	61.3	70.8	0.20
IV	57	78.3	89.2	0.22
XIII	15	70.9	92.8	0.32
XX	4	62.0	46.0	0.47
XXVI	31	77.2	85.5	0.47
XIX	6	121.3	149.3	0.48
XV	11	49.5	56.0	0.57
XXIX	5	67.2	77.6	0.60
XXVII	4	60.0	52.0	0.64
XVII	95	99.3	98.7	0.90
XVIII	2	64.0	60.0	0.91

^aAverage number of hours (h) students reported to study in preparation for a single summative in-house OSCE or MCQ test

^bNumber of responses

^cA paired t -test was used

Although other authors have reported similar findings, they only examined one OSCE at a single institution and did not use a multi-centre approach [23, 24].

In conclusion, our findings depicted that the deployment of OSCEs has an impact on the students' learning behaviour. In agreement with previous studies [23–25], the assessment tool encourages students to acquire clinical skills in, for example, physical examination,

Table 3 Performance outcomes in OSCEs and MCQ tests by medical school affiliation

Medical school	n ^b	Average grade on OSCEs ^a	Average grade on MCQ tests ^a	p ^c
		Mean	Mean	
I	22	3.52	2.90	< 0.01
V	16	3.38	2.47	< 0.01
VI	52	3.53	2.73	< 0.01
VII	42	3.61	3.00	< 0.01
IX	35	3.48	3.01	< 0.01
XI	34	3.26	2.86	< 0.01
XIV	72	3.12	2.76	< 0.01
XXI	54	2.89	2.60	< 0.01
XXIV	48	3.38	3.03	< 0.01
XXVIII	58	3.37	2.54	< 0.01
XXXI	57	3.38	2.81	< 0.01
XXVII	4	3.75	2.13	0.01
III	17	3.54	3.12	0.02
XIII	15	3.40	2.93	0.02
XXV	6	3.42	2.62	0.06
XXIX	5	3.60	3.00	0.07
XXXII	49	3.06	2.89	0.08
VIII	39	2.96	2.79	0.09
XX	4	2.50	3.00	0.09
XXII	8	3.25	3.00	0.10
XVI	43	3.24	3.08	0.13
XXX	23	2.98	2.72	0.16
XVII	105	3.05	2.95	0.19
XIX	8	3.19	2.84	0.21
X	122	2.60	2.66	0.38
XII	33	2.81	2.71	0.38
XVIII	2	3.25	2.75	0.50
XXIII	20	3.10	3.14	0.88
IV	62	2.96	2.97	0.92
XXVI	31	2.81	2.81	0.96
II	12	3.33	3.33	1.00
XV	13	2.85	3.48	0.01

^aOverall average grade across all assessments of that kind as reported by the students. We used a 4.0 grading scale (minimum–maximum, 0.0–4.0)

^bNumber of responses

^cA paired *t*-test was used

practical procedures or communication. The assessment also appears to motivate students – as compared to the MCQ tests – to focus more on studying in authentic learning environments and the community, both of which has been seen as important to support learning [26, 27].

If students prepare for an OSCE “designed to assess certain competencies or skills” [28]; vs. MCQs, which

draw items from a much larger content domain, then they would probably need less study time to achieve the required learning outcomes. Our findings confirmed this assumption for the first time. We found that even though there were differences between schools, students spent less time preparing for an OSCE compared with an MCQ test, and yet performed well.

There has been evidence that scores achieved on the OSCE are strong predictors of a later clinical performance [29, 30]. However, good performance on the OSCE does not necessarily mean a student will have the same level of competence or performance in the clinical workplace. The simulated environment in which the OSCEs take place can influence the performance. Thus, a student might perform poorer when he/she is faced with unexpected, unusual circumstances in the real workplace [31]. It is important to keep this in mind when considering the (good) performance outcomes in OSCEs [28].

Limitations

Our study has several limitations related to its sample. First, as we chose the sampling design of collecting data from every individual of the studied population by using advertisements, we could not retrace how many of the eligible students we approached. Therefore, we can report neither a response nor a non-response rate. Second, we found an overrepresentation of respondents from particular medical schools in our sample, which might have skewed the results. The varying degree to which the individual medical schools supported our study may have caused this fact. Nevertheless, the demographic profile of our sample reflected the general make-up of the medical student population in Germany and, in addition, we had a sufficiently large sample size for our analysis.

Another limitation of the study is that it relied on a self-reporting instrument to determine study resources, studying time, and performance outcomes leading to potential bias. This needs to be considered when interpreting the results.

Conclusions

We conclude that the introduction of the OSCE assessment shifted the time students spend studying. In preparation for OSCEs, students focus their attention on acquiring the necessary clinical skills, and they need less study time to achieve the expected level of competence or performance compared with MCQ tests. This clearly confirms the value of adding the OSCE assessment to a testing programme, as it places the emphasis on the acquisition of practical skills in addition to knowledge.

Additional file

Additional file 1: Questionnaire items (translated into English). (PDF 238 kb)

Abbreviations

ANOVA: Analysis of variance; MCQs: Multiple choice questions; OSCEs: Objective structured clinical examinations

Acknowledgements

The authors would like to thank the representatives of medical schools for their support, as well as Thomas Kessler from the Department of Social Psychology, University of Jena for his helpful advice on data collection and analysis. A special thanks goes to the students who completed the questionnaire.

Funding

None.

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' contributions

The study was designed and conceptualised by SM, UD, and US. SM, IK, and UD contributed to the acquisition, analysis, and interpretation of the data used for the work. SM was a major contributor in writing the manuscript and all authors read and approved the final manuscript to be published.

Ethics approval and consent to participate

According to the guidelines of the Ethics Committee of Jena University Hospital at Friedrich Schiller University, no formal ethics approval was required for this study (<https://www.uniklinikum-jena.de/ethikkommission/Antragstellung.html>). Students indicated their consent to participate by completing the questionnaire.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of General, Visceral and Vascular Surgery, Universitätsklinikum Jena, Am Klinikum 1, 07747 Jena, Germany. ²Department of Gynaecology and Reproductive Medicine, Universitätsklinikum Jena, Am Klinikum 1, 07747 Jena, Germany. ³Department of General, Visceral and Vascular Surgery, Experimental Transplantation Surgery, Universitätsklinikum Jena, Drackendorfer Str. 1, 07747 Jena, Germany.

Received: 8 August 2018 Accepted: 23 April 2019

Published online: 15 May 2019

References

1. Patrício MF, Julião M, Fareleira F, Carneiro AV. Is the OSCE a feasible tool to assess competencies in undergraduate medical education? *Med Teach*. 2013;35:503–14.
2. Boursicot K, Etheridge L, Setna Z, Sturrock A, Ker J, Smees S, Sambandam E. Performance in assessment: consensus statement and recommendations from the Ottawa conference. *Med Teach*. 2011;33:370–83.
3. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *Br Med J*. 1975;1:447–51.
4. Harden RM. Misconceptions and the OSCE. *Med Teach*. 2015;37:608–10.
5. Walters K, Osborn D, Raven P. The development, validity and reliability of a multimodality objective structured clinical examination in psychiatry. *Med Educ*. 2005;39:292–8.
6. Gilson GJ, George KE, Qualls CM, Sarto GE, Obenshain SS, Boulet J. Assessing clinical competence of medical students in women's health care: use of the objective structured clinical examination. *Obstet Gynecol*. 1998;92:1038–43.
7. Matsell DG, Wolfish NM, Hsu E. Reliability and validity of the objective structured clinical examination in paediatrics. *Med Educ*. 1991;25:293–9.
8. Petrusa ER, Blackwell TA, Rogers LP, Saydjari C, Parcel S, Guckian JC. An objective measure of clinical performance. *Am J Med*. 1987;83:34–42.
9. Pierre RB, Wierenga A, Barton M, Branday JM, Christie CDC. Student evaluation of an OSCE in paediatrics at the University of the West Indies, Jamaica. *BMC Med Educ*. 2004;4:22.
10. Wilkinson TJ, Newble DI, Wilson PD, Carter JM, Helms RM. Development of a three-centre simultaneous objective structured clinical examination. *Med Educ*. 2000;34:798–807.
11. Newble DI. Eight years' experience with a structured clinical examination. *Med Educ*. 1988;22:200–4.
12. Newble DI, Jaeger K. The effect of assessments and examinations on the learning of medical students. *Med Educ*. 1983;17:165–71.
13. Mavis BE. Does studying for an objective structured clinical examination make a difference? *Med Educ*. 2000;34:808–12.
14. Rudland J, Wilkinson T, Smith-Han K, Thompson-Fawcett M. "You can do it late at night or in the morning. You can do it at home, I did it with my flatmate." The educational impact of an OSCE. *Med Teach*. 2008;30:206–11.
15. Hoffman KG, Donaldson JF. Contextual tensions of the clinical environment and their influence on teaching and learning. *Med Educ*. 2004;38:448–54.
16. Approbationsordnung für Ärzte [National medical licensure act]. 2002. https://www.gesetze-im-internet.de/_appro_2002/BJNR240500002.html. Accessed 30 May 2018.
17. Müller S, Dahmen U, Settmacher U. Application of the objective structured clinical examination (OSCE) in German medical schools: an inventory. *Gesundheitswesen*. 2018;80:1099–103.
18. Destatis: Bildung und Kultur – Studierende an Hochschulen [Education and culture – students in higher education]. https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/DEHeft_derivate_00016700/2110410157004.pdf. Accessed 29 Apr 2019.
19. 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:Doc51.
20. Broekkamp H, van Hout-Wolters BHAM. Students' adaptation of study strategies when preparing for classroom tests. *Educ Psychol Rev*. 2007;19:401–28.
21. Biggs J. Assessing learning quality: reconciling institutional, staff and educational demands. *Assess Eval High Educ*. 1996;21:5–16.
22. Cilliers FJ, Schuwirth LWT, van der Vleuten CPM. A model of the pre-assessment learning effects of assessment is operational in an undergraduate clinical context. *BMC Med Educ*. 2012;12:9.
23. Kowlowitz V, Hoole AJ, Sloane PD. Implementing the objective structured clinical examination in a traditional medical school. *Acad Med*. 1991;66:345–7.
24. Johnson G, Reynard K. Assessment of an objective structured clinical examination (OSCE) for undergraduate students in accident and emergency medicine. *J Accid Emerg Med*. 1994;11:223–6.
25. Duvivier RJ, van Geel K, van Dalen J, Scherpier AJJA, van der Vleuten CPM. Learning physical examination skills outside timetabled training sessions: what happens and why? *Adv Health Sci Educ Theory Pract*. 2012;17:339–55.
26. Durning SJ, Artino AR. Situativity theory: a perspective on how participants and the environment can interact: AMEE guide no. 52. *Med Teach*. 2011;33:188–99.
27. Mann KV. Theoretical perspectives in medical education: past experience and future possibilities. *Med Educ*. 2011;45:60–8.
28. Khan KZ, Ramachandran S, Gaunt K, Pushkar P. The objective structured clinical examination (OSCE): AMEE guide no. 81. Part I: an historical and theoretical perspective. *Med Teach*. 2013;35:e1437–46.
29. Wallenstein J, Heron S, Santen S, Shayne P, Ander D. A core competency-based objective structured clinical examination (OSCE) can predict future resident performance. *Acad Emerg Med*. 2010;17(Suppl 2):67–71.

30. Probert CS, Cahill DJ, McCann GL, Ben-Shlomo Y. Traditional finals and OSCEs in predicting consultant and self-reported clinical skills of PRHOs: a pilot study. *Med Educ.* 2003;37:597–602.
31. Khan K, Ramachandran S. Conceptual framework for performance assessment: competency, competence and performance in the context of assessments in healthcare – deciphering the terminology. *Med Teach.* 2012;34:920–8.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



Objective Structured Clinical Examination (OSCE)

What does the OSCE do for you?

	<i>strongly disagree</i>	<i>disagree</i>	<i>neither agree nor disagree</i>	<i>agree</i>	<i>strongly agree</i>
OSCE gives me an understanding of medical care	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
demonstrates the practices and principles of medical treatment	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
gives me feedback on my performance level	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
reveals my strengths in medical practice	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
reveals my weaknesses in medical practice	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
shows me gaps in my education	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
enhances my problem-solving and decision-making abilities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
promotes my theoretical knowledge	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
reflects the requirements of the medical profession	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
allows me to assess my own ability to work as a medical professional	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
helps me with my speciality choice	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Objective Structured Clinical Examination (OSCE)

How important are each of the following resources to prepare for OSCEs?

	<i>not important</i>	<i>slightly important</i>	<i>moderately important</i>	<i>important</i>	<i>very important</i>
Physical examination courses	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Skills lab	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Clinical work placements	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Medical clerkships	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Peer tutorials	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Group learning	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
PBL courses	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Multimedia materials	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Lectures	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Textbooks	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Casebooks	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Class notes/logs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Objective Structured Clinical Examination (OSCE)

How much time do you spend preparing for an OSCE?

Average total preparation time for **one** summative OSCE, including course attendance

days

(one day = approx. 8 working hours)

Objective Structured Clinical Examination (OSCE)

What are your OSCE results?

Your average OSCE result

[grade]

Multiple Choice Questions (MCQs)

What does the MCQ format do for you?

	<i>strongly disagree</i>	<i>disagree</i>	<i>neither agree nor disagree</i>	<i>agree</i>	<i>strongly agree</i>
MCQ format gives me an understanding of medical care	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
demonstrates the practices and principles of medical treatment	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
gives me feedback on my performance level	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
reveals my strengths in medical practice	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
reveals my weaknesses in medical practice	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
shows me gaps in my education	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
enhances my problem-solving and decision-making abilities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
promotes my theoretical knowledge	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
reflects the requirements of the medical profession	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
allows me to assess my own ability to work as a medical professional	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
helps me with my speciality choice	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Multiple Choice Questions (MCQs)

How important are each of the following resources to prepare for MCQ tests?

	<i>not important</i>	<i>slightly important</i>	<i>moderately important</i>	<i>important</i>	<i>very important</i>
Physical examination courses	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Skills lab	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Clinical work placements	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Medical clerkships	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Peer tutorials	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Group learning	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
PBL courses	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Multimedia materials	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Lectures	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Textbooks	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Casebooks	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Class notes/logs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Multiple Choice Questions (MCQs)

How much time do you spend preparing for an MCQ test?

Average total preparation time for **one** summative MCQ test, including course attendance

days

(one day = approx. 8 working hours)

Multiple Choice Questions (MCQs)

What are your MCQ test results?

Your average MCQ test result

[grade]

Demographics

What is your gender?

Female **Male**

How old are you?

years

Which semester are you in?

[Please select] ▼

What medical school are you from?

[Please select] ▼

- 4.3** *A pilot survey of student perceptions on the benefit of the OSCE and MCQ modalities, Stefan Müller, Utz Settmacher, Ines Koch, Uta Dahmen, GMS J Med Educ, 35(4):Doc51, 2018*

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											<i>n</i>
	A	B	C	D	E	F	G	H	J	K	L	
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 betragen 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

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

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

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 Novellierung 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	<i>p</i>
	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											<i>n</i>
	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 niedriger 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, 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

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/>.

5 Diskussion

Diese Arbeit hat einen genaueren Blick auf die OSCE-Prüfungsmethode in der medizinischen Ausbildung der Bundesrepublik Deutschland geworfen. Die einzelnen Studien wurden initiiert, um erstens nähere Informationen über den Einsatz von OSCEs an den medizinischen Fakultäten zu erhalten, zweitens herauszufinden, wie die OSCEs das Lernen der Studierenden beeinflussen, und drittens Datenmaterial zu gewinnen, das eine Einbeziehung des OSCE in das medizinische Staatsexamen unterstützt. Es konnten bundesweit Daten erhoben werden. Die wichtigsten Ergebnisse der Studien waren wie folgt: Fast jede Fakultät hat zumindest einen OSCE im Curriculum implementiert, die OSCEs beeinflussen das Lernen der Studierenden in der gewünschten Weise, und die Studierenden betrachten den OSCE als nützliches Prüfungsformat.

Prüfungen im Medizinstudium waren lange Zeit überwiegend auf den Abruf von theoretischem Fachwissen durch MC-Prüfungsaufgaben ausgelegt. Eine Überprüfung praktischer Fähigkeiten und Fertigkeiten erfolgte dagegen kaum (Falck-Ytter und Stiegler 1993, Schulze et al. 2004). Durch die Reform der ÄApprO im Jahr 2002 waren die Fakultäten aufgefordert, ihre Prüfungsstrukturen zu verändern und Formate einzusetzen, die der Vorgabe nach mehr Praxisorientierung gerecht werden. In der vorliegenden Abhandlung konnte nachgewiesen werden, dass OSCE-Prüfungen zwischenzeitlich weitverbreitet sind. Obgleich nahezu alle Fakultäten berichteten, in irgendeiner Form die OSCE-Methode zu verwenden, variiert das Ausmaß des Einsatzes erheblich zwischen den Ausbildungsstätten. Auch in Bezug auf die Inhalte und Ziele der OSCEs gibt es starke Unterschiede. Es lässt sich resümieren, dass die Umstellung auf praxisnahe Prüfungen an den deutschen Fakultäten trotz großer Fortschritte noch nicht hinreichend vollzogen ist.

Für die Ausübung des Arztberufs sind Fähigkeiten bzw. Fertigkeiten in ärztlicher Gesprächsführung, körperlicher Untersuchung oder der Durchführung praktischer Maßnahmen von grundlegender Bedeutung (Holmboe 2004, Sawyer et al. 2015). Die Ergebnisse der Arbeit haben bestätigt, dass die Studierenden gemäß dem Prinzip „assessment drives learning“ durch OSCE-Prüfungen ermuntert werden, sich Kompetenzen in diesen Bereichen anzueignen. Dies unterstreicht den Mehrwert, den die OSCE-Methode für ein Prüfungssystem in der medizinischen Ausbildung hat.

Das OSCE-Format ist bislang nicht Bestandteil der deutschen Staatsexamensprüfungen (*Kapitel 2.4*). Der Masterplan Medizinstudium 2020, der von der Bundesregierung und den Ländern zur Reformierung des Medizinstudiums beschlossen wurde, sieht jedoch eine Neugestaltung der Staatsexamina vor. Zukünftig sollen die Studierenden am zweiten Tag der dritten staatlichen Prüfung mithilfe eines OSCE geprüft werden. Um bundesweit einheitliche Standards bei den OSCE-Prüfungen zu schaffen, sollen die Prüfungsinhalte sowie die Beurteilungskriterien für alle Fakultäten verbindlich durch das IMPP vorgegeben werden (Bundesministerium für Bildung und Forschung 2017, Jünger 2018). Die erhobenen Daten haben gezeigt, dass die Studierenden den OSCE als wirkungsvolle Lernerfahrung und als hilfreich für die Entwicklung klinischer bzw. praktischer Fähigkeiten sehen. Das sind weitere Argumente, die für eine Einbeziehung des OSCE in die deutschen Staatsprüfungen sprechen.

5.1 Kompetenzorientierte medizinische Ausbildung

Der Trend in der medizinischen Ausbildung geht von einer struktur- und prozessbasierten zu einer kompetenzorientierten Ausbildung (Carraccio et al. 2002, Hawkins et al. 2015). Das Konzept der kompetenzorientierten medizinischen Ausbildung basiert auf einem nach Kompetenzen ausgerichteten Bildungsmodell, das sich im Wesentlichen an dem vorgegebenen Absolventenprofil orientiert. Lehrplan und Prüfungen werden um die definierten Standards der Absolventen-Outcomes (Kompetenzen) herum organisiert und sind aufeinander abgestimmt (Frank et al. 2010a, 2010b). Für die erfolgreiche Umsetzung des Konzepts ist eine Implementierung umfassender und ausgefeilter Prüfungssysteme mit Schwerpunktsetzung auf die Anforderungen an die ärztliche Tätigkeit erforderlich. Somit werden performancebasierte Prüfungen in Zukunft noch weiter an Bedeutung gewinnen (Holmboe et al. 2010).

In Deutschland wurde auf Grundlage des CanMEDS-Rahmenwerks (Frank und Danoff 2007) zu den sieben Rollen des Arztes (medizinischer Experte, Kommunikator, Mitglied eines Teams, Verantwortungsträger und Manager, Gesundheitsberater und -fürsprecher, Gelehrter, professionell Handelnder) der Nationale Kompetenzbasierte Lernzielkatalog Medizin (NKLM) erarbeitet. Der NKLM beschreibt als eine Art Kerncurriculum die Kompetenzen, die Medizinstudierende am Ende des Studiums zur Befähigung für die ärztliche Weiterbildung haben sollten. Der Katalog ist bislang nicht verbindlich

und soll den Fakultäten als Orientierung für die Weiterentwicklung ihrer Curricula dienen (Medizinischer Fakultätentag 2015, Fischer et al. 2015).

Mit der Erarbeitung des NKLM und dessen Verabschiedung im Juni 2015 (Fischer et al. 2015) sowie dem Masterplan, der eine Kompetenzorientierung im Studium und den Staatsexamina vorsieht (Bundesministerium für Bildung und Forschung 2017, Jünger 2018), wurden in Deutschland die ersten Schritte in Richtung einer kompetenzorientierten medizinischen Ausbildung unternommen. Die Fakultäten sind nun am Zug, neben der kompetenzorientierten Ausgestaltung ihrer Curricula den Wandel in der Ausbildung auch im Prüfungswesen zu vollziehen. Dafür wird ein Ausbau des Einsatzes von OSCE-Prüfungen sowie eine breitere inhaltliche Abdeckung im Rahmen dieser Prüfungen unausweichlich sein.

5.2 Probleme und Herausforderungen

Eine Ausweitung des fakultätsinternen Einsatzes von OSCEs bzw. eine Einbeziehung des OSCE in die Staatsexamensprüfungen hat jedoch ihre Herausforderungen, die es zu bewältigen gilt. Die Entwicklung und Durchführung von OSCE-Prüfungen ist sehr arbeits- und personalintensiv (Boursicot et al. 2010). Durch eine gute Organisation und den bedachten Einsatz vorhandener Ressourcen können jedoch Kosten eingespart werden. Vorgeschlagen wird z. B. eine gemeinschaftliche Nutzung von Stationen zwischen Fakultäten, das Einsetzen studentischer Hilfskräfte als Stationsentwickler und Prüfer oder die Verwendung echter Patienten anstelle von Schauspielpatienten. Wenngleich das Ziel bzw. die Art der Prüfung (summativ vs. formativ, Staatsexamen vs. fakultätsinterne Prüfung) die Ausgaben durch Unterschiede in der Prüfungssicherheit und der aufzuwendenden Zeit beeinflusst, ist die Implementierung eines OSCE auch in Einrichtungen mit kleinerem Budget möglich (Poenaru et al. 1997, Casey et al. 2009, Harden et al. 2016).

Als ein anderes Problem an den deutschen Fakultäten kann die geringe Erfahrung mit der OSCE-Prüfungsmethode gesehen werden. Einen OSCE von Grund auf neu zu planen und zu organisieren, ist eine gewaltige Aufgabe, die umfangreiches Wissen, Erfahrung und Schulung für alle Beteiligten erfordert. Wenn Mitarbeiter der Fakultät bereits über Erfahrung oder Expertenwissen in der Ausrichtung von OSCE-Prüfungen verfügen, sollten diese Personen mit der Leitung der Prüfung beauftragt werden (Khan et al. 2013). Ist dies nicht der Fall, bietet die Literatur ausführliche Informationen, die den Fakultäten bei

der effektiven Umsetzung eines OSCE-Programms helfen sollen (Hodges et al. 2002, Khan et al. 2013, Harden et al. 2016). Gleichzeitig können Graduiertenprogramme wie der Studiengang Master of Medical Education (MME) Deutschland einen wichtigen Beitrag zur entsprechenden Qualifizierung des Lehrpersonals leisten (Jünger et al. 2008).

Weitere Probleme im Zusammenhang mit einem Ausbau des Einsatzes von OSCE-Prüfungen dürften wie andernorts fehlende Zeit und mangelndes Interesse seitens der Dozierenden sowie immanente konservative Strukturen an den Fakultäten sein (Fowell et al. 2000, Lee und Ahn 2006). Um diese Probleme zu lösen, sollten zunächst Anreize für ein größeres Engagement des ärztlichen Personals im Bereich der Lehre geschaffen werden. Graduiertenprogramme scheinen auch hier vielversprechende Instrumente zu sein (Sethi et al. 2016). Zugleich sollte das Lehrpersonal in höherem Maße von konkurrierenden Verpflichtungen in Forschung und Patientenversorgung entbunden werden. Neben der angesprochenen Aufwertung der Lehrtätigkeit bedarf es an den Fakultäten darüber hinaus einer noch stärkeren Bereitschaft zur Veränderung etablierter Systeme. Die Errichtung eines Zentrums für medizinische Ausbildung innerhalb der Fakultät kann hierzu entscheidende Impulse geben (Davis et al. 2005).

5.3 Stärken und Schwächen der Arbeit

Keine einzige Prüfungsmethode ist imstande, sämtliche Aspekte klinischer Kompetenz adäquat zu erfassen. Eine Kombination verschiedener Prüfungsverfahren gilt daher als der beste Ansatz, um die klinische Kompetenz einer Person ganzheitlich beurteilen zu können (Schuwirth und van der Vleuten 2010). Die vorliegende Arbeit hat sich mit dem OSCE als der zentralen Form einer performancebasierten Prüfung beschäftigt. Anzumerken ist, dass arbeitsplatzbezogene Prüfungen wie Mini-CEX (Clinical Evaluation Exercise) oder DOPS (Direct Observation of Procedural Skills), die am echten klinischen Arbeitsplatz stattfinden (Norcini und McKinley 2007) und entsprechende Formate für die oberste Stufe der „Miller-Pyramide“ sind (*Kapitel 2.2*), nicht Gegenstand der durchgeführten Studien waren.

In den Studien wurden verschiedene Methoden der Umfrageforschung verwendet. Umfragen sind ein gängiges Mittel in der Ausbildungsforschung von Gesundheitsberufen (Artino et al. 2014, Phillips et al. 2017). Umfragen haben den Vorteil, dass in relativ kurzer Zeit und mit geringem finanziellen Aufwand eine große Menge empirischer Daten

erhoben werden kann (Kelley et al. 2003). Für die Durchführung guter Umfragen bedarf es wie bei anderen Forschungen einer soliden Planung und Vorbereitung. Dazu gehört die Festlegung eines klaren Forschungsziels, die Entwicklung passender Fragen sowie die Wahl einer geeigneten Forschungsmethode (Jones et al. 2013).

Für die Bestandsaufnahme zum Einsatz des OSCE an den Fakultäten (Studie I) wurde die Forschungsmethode der Telefoninterviews bzw. E-Mail-Korrespondenz gewählt. Eine Stärke der semistrukturierten Interviews (Fragenkatalog mit individuell gestalteter Abfolge der Fragen) bestand in dem wechselseitigen Dialog zwischen Befrager und Befragtem. Dadurch konnten komplizierte Sachverhalte geklärt und teilweise noch zusätzliche Informationen gewonnen werden. Zugleich führte der Interviewansatz zu einer sehr hohen Rücklaufquote, die durch die postalische oder elektronische Versendung eines Fragebogens wohl kaum zu erreichen gewesen wäre (Kelley et al. 2003, Jones et al. 2013).

Die Forschungsmethode des Fragebogens eignet sich besonders gut zur Erhebung von Daten über Meinungen, Einstellungen und Überzeugungen. Darüber hinaus kann ein Fragebogen nützlich sein, um Informationen zu Verhaltensweisen zu bekommen, die nicht direkt beobachtbar sind, wie z. B. Lernstrategien (Artino et al. 2014). Zur Erfassung des Lernverhaltens der Studierenden im Vorfeld von OSCEs (Studie II) und der studentischen Wahrnehmung zum Nutzen des OSCE (Studie III) wurde ein Fragebogen konzipiert. Die Forschungsmethode hatte allerdings ihre Schwächen: Im Gegensatz zu den Interviews erlaubte das Konzept des Fragebogens keine unmittelbare Interaktion zwischen Befrager und Befragtem. Obgleich durch einen umfangreichen Pretest schwer- oder missverständliche bzw. unterschiedlich interpretierbare Fragen identifiziert und nachfolgend modifiziert wurden (Krosnick 1999), könnte es dennoch zu Antwortverzerrungen durch Fehlinterpretationen gekommen sein. Gleichzeitig erforderte die Forschungsmethode eine ausreichend große Stichprobe, um zu gewährleisten, dass die Ergebnisse der Befragung möglichst repräsentativ für die Gesamtpopulation sind (Kelley et al. 2003).

6 Schlussfolgerungen

Im Masterplan wurde festgeschrieben, einen OSCE in die Staatsexamensprüfungen einzubeziehen. Die vorgestellte Arbeit hat den aktuellen Einsatz der OSCE-Methode in Deutschland beleuchtet. Zum einen stellte sich heraus, dass OSCE-Prüfungen an den Fakultäten bei Weitem noch nicht überall dort eingesetzt werden, wo es möglich oder angebracht wäre. Zum anderen konnte festgestellt werden, dass OSCEs einen positiven Effekt auf das studentische Lernverhalten haben und von den Studierenden als sehr sinnvoll wahrgenommen werden. Die Fakultäten sollten den Einsatz der Prüfungsmethode daher ausbauen und sicherstellen, dass zumindest in den examensrelevanten Fächern OSCE-Prüfungen durchgeführt werden.

7 Literatur- und Quellenverzeichnis

- Ainsworth MA, Callaway MR, Perkowski L. 1995. An OSCE assessment of fourth-year students. *Acad Med*, 70(5):444–445.
- Approbationsordnung für Ärzte vom 27. Juni 2002 (BGBl. I S. 2405), die zuletzt durch Artikel 5 des Gesetzes vom 17. Juli 2017 (BGBl. I S. 2581) geändert worden ist.
- Artino AR Jr, La Rochelle JS, Dezee KJ, Gehlbach H. 2014. Developing questionnaires for educational research: AMEE Guide No. 87. *Med Teach*, 36(6):463–474.
- Boulet JR, Smee SM, Dillon GF, Gimpel JR. 2009. The use of standardized patient assessments for certification and licensure decisions. *Simul Healthc*, 4(1):35–42.
- Boursicot KAM, Roberts TE, Burdick WP. 2010. Structured assessments of clinical competence. In: Swanwick T, Hrsg. *Understanding Medical Education: Evidence, Theory and Practice*. Chichester: Wiley, 246–258.
- Bundesministerium für Bildung und Forschung (BMBF). 2017. Masterplan Medizinstudium 2020. URL: https://www.bmbf.de/files/2017-03-31_Masterplan%20Beschlusstext.pdf (zugegriffen am 7. Dezember 2017).
- Carraccio C, Wolfsthal SD, Englander R, Ferentz K, Martin C. 2002. Shifting paradigms: from Flexner to competencies. *Acad Med*, 77(5):361–367.
- Casey PM, Goepfert AR, Espey EL, Hammoud MM, Kaczmarczyk JM, Katz NT, Neutens JJ, Nuthalapaty FS, Peskin E. 2009. To the point: reviews in medical education – the Objective Structured Clinical Examination. *Am J Obstet Gynecol*, 200(1):25–34.
- Cilliers FJ, Schuwirth LW, Adendorff HJ, Herman N, van der Vleuten CP. 2010. The mechanism of impact of summative assessment on medical students' learning. *Adv Health Sci Educ Theory Pract*, 15(5):695–715.
- Cilliers FJ, Schuwirth LWT, Herman N, Adendorff HJ, van der Vleuten CPM. 2012. A model of the pre-assessment learning effects of summative assessment in medical education. *Adv Health Sci Educ Theory Pract*, 17(1):39–53.
- Cook DA, Brydges R, Ginsburg S, Hatala R. 2015. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ*, 49(6):560–575.
- Davis MH. 2003. OSCE: the Dundee experience. *Med Teach*, 25(3):255–261.

- Davis MH, Karunathilake I, Harden RM. 2005. AMEE Education Guide no. 28: The development and role of departments of medical education. *Med Teach*, 27(8):665–675.
- Downing SM. 2003. Validity: on the meaningful interpretation of assessment data. *Med Educ*, 37(9):830–837.
- Epstein RM. 2007. Assessment in medical education. *N Engl J Med*, 356(4):387–396.
- Epstein RM, Hundert EM. 2002. Defining and assessing professional competence. *JAMA*, 287(2):226–235.
- Falck-Ytter Y, Stiegler I. 1993. Beurteilung klinischer Fähigkeiten mit dem OSCE-Verfahren. *Med Ausbild*, 10(1):48–55.
- Fernandez N, Dory V, Ste-Marie LG, Chaput M, Charlin B, Boucher A. 2012. Varying conceptions of competence: an analysis of how health sciences educators define competence. *Med Educ*, 46(4):357–365.
- Fischer MR, Bauer D, Mohn K. 2015. Endlich fertig! Nationale Kompetenzbasierte Lernzielkataloge Medizin (NKLM) und Zahnmedizin (NKLZ) gehen in die Erprobung. *GMS Z Med Ausbild*, 32(3):Doc35.
- Fowell SL, Maudsley G, Maguire P, Leinster SJ, Bligh J. 2000. Student assessment in undergraduate medical education in the United Kingdom, 1998. *Med Educ*, 34 Suppl:S1–S49.
- Frank JR, Danoff D. 2007. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. *Med Teach*, 29(7):642–647.
- Frank JR, Mungroo R, Ahmad Y, Wang M, De Rossi S, Horsley T. 2010a. Toward a definition of competency-based education in medicine: a systematic review of published definitions. *Med Teach*, 32(8):631–637.
- Frank JR, Snell LS, Ten Cate O, Holmboe ES, Carraccio C, Swing SR, Harris P, Glasgow NJ, Campbell C, Dath D, Harden RM, Iobst W, Long DM, Mungroo R, Richardson DL, Sherbino J, Silver I, Taber S, Talbot M, Harris KA. 2010b. Competency-based medical education: theory to practice. *Med Teach*, 32(8):638–645.
- Gross M, Pelz J. 2009. Veränderung des Berufsbildes des Arztes. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*, 52(8):831–840.
- Haage H. 2002. Das neue Medizinstudium. *Medizinrecht*, 20(9):456–461.

- Harden RM. 2006. Trends and the future of postgraduate medical education. *Emerg Med J*, 23(10):798–802.
- Harden RM, Lilley P, Patrício M. 2016. *The Definitive Guide to the OSCE: The Objective Structured Clinical Examination as a performance assessment*. Edinburgh: Elsevier.
- Harden RM, Stevenson M, Downie WW, Wilson GM. 1975. Assessment of clinical competence using objective structured examination. *Br Med J*, 1(5955):447–451.
- Härtl A, Berberat P, Fischer MR, Forst H, Grützner S, Händl T, Joachimski F, Linné R, Märkl B, Naumann M, Putz R, Schneider W, Schöler C, Wehler M, Hoffmann R. 2017. Development of the competency-based medical curriculum for the new Augsburg University Medical School. *GMS J Med Educ*, 34(2):Doc21.
- Hawkins RE, Welcher CM, Holmboe ES, Kirk LM, Norcini JJ, Simons KB, Skochelak SE. 2015. Implementation of competency-based medical education: are we addressing the concerns and challenges? *Med Educ*, 49(11):1086–1102.
- Heeneman S, Oudkerk Pool A, Schuwirth LWT, van der Vleuten CPM, Driessen EW. 2015. The impact of programmatic assessment on student learning: theory versus practice. *Med Educ*, 49(5):487–498.
- Hodges B. 2003. OSCE! Variations on a theme by Harden. *Med Educ*, 37(12):1134–1140.
- Hodges B, Hanson M, McNaughton N, Regehr G. 2002. Creating, monitoring, and improving a psychiatry OSCE: a guide for faculty. *Acad Psychiatry*, 26(3):134–161.
- Holmboe ES. 2004. Faculty and the observation of trainees' clinical skills: problems and opportunities. *Acad Med*, 79(1):16–22.
- Holmboe ES, Sherbino J, Long DM, Swing SR, Frank JR. 2010. The role of assessment in competency-based medical education. *Med Teach*, 32(8):676–682.
- Huber-Lang M, Palmer A, Grab C, Boeckers A, Boeckers TM, Oechsner W. 2017. Visions and reality: the idea of competence-oriented assessment for German medical students is not yet realised in licensing examinations. *GMS J Med Educ*, 34(2):Doc25.
- Jones TL, Baxter MAJ, Khanduja V. 2013. A quick guide to survey research. *Ann R Coll Surg Engl*, 95(1):5–7.

- Jünger J. 2018. Kompetenzorientiert prüfen im Staatsexamen Medizin. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz, 61(2):171–177.
- Jünger J, Fischer MR, Duelli R, Putz R, Resch F. 2008. Konzeption, Implementierung und Evaluation eines interfakultären Master of Medical Education Programms. Z Evid Fortbild Qual Gesundheitswes, 102(10):620–627.
- Kelley K, Clark B, Brown V, Sitzia J. 2003. Good practice in the conduct and reporting of survey research. Int J Qual Health Care, 15(3):261–266.
- Khan KZ, Gaunt K, Ramachandran S, Pushkar P. 2013. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part II: organisation & administration. Med Teach, 35(9):e1447–e1463.
- Kim KS. 2010. Introduction and administration of the clinical skill test of the medical licensing examination, Republic of Korea (2009). J Educ Eval Health Prof, 7:4.
- Kollewe T, Sennekamp M, Ochsendorf F. 2018. Medizindidaktik – Erfolgreich lehren und Wissen vermitteln. Berlin: Springer, 97–124.
- Krosnick JA. 1999. Survey research. Annu Rev Psychol, 50:537–567.
- Lee M, Wimmers PF. 2011. Clinical competence understood through the construct validity of three clerkship assessments. Med Educ, 45(8):849–857.
- Lee YM, Ahn DS. 2006. The OSCE: a new challenge to the evaluation system in Korea. Med Teach, 28(4):377–379.
- McNaughton N, Ravitz P, Wadell A, Hodges BD. 2008. Psychiatric education and simulation: a review of the literature. Can J Psychiatry, 53(2):85–93.
- Medizinischer Fakultätentag (MFT), Hrsg. 2015. Nationaler Kompetenzbasierter Lernzielkatalog Medizin. URL: http://www.nklm.de/files/nklm_final_2015-07-03.pdf (zugegriffen am 2. November 2017).
- Miller GE. 1990. The assessment of clinical skills/competence/performance. Acad Med, 65(9):S63–S67.
- Newble D. 2016. Revisiting ‘The effect of assessments and examinations on the learning of medical students’. Med Educ, 50(5):498–501.
- Nikendei C, Krautter M, Celebi N, Obertacke U, Jünger J. 2012. Final year medical education in Germany. Z Evid Fortbild Qual Gesundheitswes, 106(2):75–84.
- Nikendei C, Weyrich P, Jünger J, Schrauth M. 2009. Medical education in Germany. Med Teach, 31(7):591–600.

- Norcini JJ, McKinley DW. 2007. Assessment methods in medical education. *Teach Educ*, 23(3):239–250.
- Petrusa ER. 2002. Clinical performance assessments. In: Norman GR, van der Vleuten CPM, Newble DI, Hrsg. *International Handbook of Research in Medical Education*. Dordrecht: Kluwer Academic Publishers, 673–709.
- Phillips AW, Friedman BT, Utrankar A, Ta AQ, Reddy ST, Durning SJ. 2017. Surveys of health professions trainees: prevalence, response rates, and predictive factors to guide researchers. *Acad Med*, 92(2):222–228.
- Poenaru D, Morales D, Richards A, O'Connor HM. 1997. Running an objective structured clinical examination on a shoestring budget. *Am J Surg*, 173(6):538–541.
- Reznick RK, Blackmore D, Dauphinée WD, Rothman AI, Smee S. 1996. Large-scale high-stakes testing with an OSCE: report from the Medical Council of Canada. *Acad Med*, 71(1):S19–S21.
- Roberts C, Newble D, Jolly B, Reed M, Hampton K. 2006. Assuring the quality of high-stakes undergraduate assessments of clinical competence. *Med Teach*, 28(6):535–543.
- Rust C. 2002. The impact of assessment on student learning: How can the research literature practically help to inform the development of departmental assessment strategies and learner-centred assessment practices? *Act Learn High Educ*, 3(2):145–158.
- Sadler DR. 1989. Formative assessment and the design of instructional systems. *Instr Sci*, 18(2):119–144.
- Sawyer T, White M, Zaveri P, Chang T, Ades A, French H, Anderson JD, Auerbach M, Johnston L, Kessler D. 2015. Learn, see, practice, prove, do, maintain: An evidence-based pedagogical framework for procedural skill training in medicine. *Acad Med*, 90(8):1025–1033.
- Schulze J, Drolshagen S, Nürnberger F, Siegers CP, Syed Ali S. 2004. Prüfen und Prüfungen nach der neuen Approbationsordnung – Grundsätze und Rahmenbedingungen. *Med Ausbildung*, 21(1):30–34.
- Schuwirth LWT, van der Vleuten CPM. 2010. How to design a useful test: the principles of assessment. In: Swanwick T, Hrsg. *Understanding Medical Education: Evidence, Theory and Practice*. Chichester: Wiley, 195–207.

- Schuwirth LWT, van der Vleuten CPM. 2011. Programmatic assessment: From assessment of learning to assessment for learning. *Med Teach*, 33(6):478–485.
- Sethi A, Schofield S, Ajjawi R, McAleer S. 2016. How do postgraduate qualifications in medical education impact on health professionals? *Med Teach*, 38(2):162–167.
- Smith LJ, Price DA, Houston IB. 1984. Objective structured clinical examination compared with other forms of student assessment. *Arch Dis Child*, 59(12):1173–1176.
- Taras M. 2005. Assessment – summative and formative – some theoretical reflections. *Br J Educ Stud*, 53(4):466–478.
- Townsend AH, McIlvenny S, Miller CJ, Dunn EV. 2001. The use of an objective structured clinical examination (OSCE) for formative and summative assessment in a general practice clinical attachment and its relationship to final medical school examination performance. *Med Educ*, 35(9):841–846.
- Via medici. 2017. Auf einen Blick: Alle Studienorte und ihre Kriterien. URL: <https://www.thieme.de/viamedici/vor-dem-studium-infos-zum-medizinstudium-1493/a/alle-studienorte-medizin-und-ihre-kriterien-33843.htm> (zugegriffen am 20. November 2017).
- Wood DF. 2013. Formative assessment. In: Walsh K, Hrsg. *Oxford Textbook of Medical Education*. Oxford: Oxford University Press, 478–488.
- Wood T. 2009. Assessment not only drives learning, it may also help learning. *Med Educ*, 43(1):5–6.
- Zavlin D, Jubbal KT, Noé JG, Gansbacher B. 2017. A comparison of medical education in Germany and the United States: from applying to medical school to the beginnings of residency. *Ger Med Sci*, 15:Doc15.

8 Anhang

Danksagung

Mein Dank gilt zuerst Frau Prof. Dr. Uta Dahmen für die ausgezeichnete wissenschaftliche Betreuung und dass sie es mir ermöglicht hat, im Bereich Ausbildungsforschung zu promovieren.

Zugleich möchte ich mich bei den weiteren Co-Autoren – Herrn Prof. Dr. Utz Settmacher und Frau Dr. Ines Koch – für die gute Zusammenarbeit bedanken.

Darüber hinaus danke ich Frau Dr. Heike Hoyer und ganz besonders Herrn Prof. Dr. Rolf Steyer und Herrn Prof. Dr. Thomas Kessler für ihren Einsatz und die wertvollen Anregungen zur Datenanalyse.

Abschließend möchte ich mich noch bei meiner Familie für die uneingeschränkte Unterstützung während der Promotionszeit ganz herzlich bedanken.

Lebenslauf

Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass mir die Promotionsordnung der Medizinischen Fakultät der Friedrich-Schiller-Universität bekannt ist,

ich die Dissertation selbst angefertigt habe und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in meiner Arbeit angegeben sind,

mich folgende Personen bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts unterstützt haben: Prof. Dr. Uta Dahmen, Prof. Dr. Utz Settmacher, Dr. Ines Koch, Prof. Dr. Rolf Steyer, Prof. Dr. Thomas Kessler und Dr. Heike Hoyer,

die Hilfe eines Promotionsberaters nicht in Anspruch genommen wurde und dass Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen,

dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe und

dass ich die gleiche, eine in wesentlichen Teilen ähnliche oder eine andere Abhandlung nicht bei einer anderen Hochschule als Dissertation eingereicht habe.

Jena, den 24. Juni 2019

Stefan Müller