

Die Freiburger Performanzprüfung – ein ressourcensparendes Prüfungsformat zur körperlichen Untersuchung

Performance Test – Low Cost Exam Format for Physical Check-Up

Lena Krabbe¹, Irmgard Streitlein-Böhme², August Sigle¹, Andreas Möltner³, Wilhelm Niebling¹, Klaus Böhme²

Hintergrund

Eine kompetenzorientierte und praxisnahe Ausbildung von Medizinstudierenden setzt – entsprechend dem *Constructive Alignment-Modell* – praxisnahe Prüfungen voraus. Die Medizinstudierenden der Universität Freiburg schließen einen mehrstufigen Kurs zur Vermittlung von Basistechniken zur körperlichen Untersuchung mit der „Famulatureifepfung“, einer sog. Performanzprüfung, ab. Dabei müssen sie in einem von sechs möglichen, zufällig zugelosten Prüfungsmodul ihre Fertigkeit zur körperlichen Untersuchung unter Beweis stellen. Es soll betrachtet werden, ob diese Vorgehensweise im Hinblick auf die Reliabilität der Prüfung vertretbar ist.

Methoden

Die Prüfungsergebnisse der 73 Studierenden, die im Sommersemester 2018 diese Famulatureifepfung ablegten, wurden nach einer deskriptiven Auswertung einer Mixed-Model-Varianzanalyse sowie einer Berechnung der Reliabilität unterzogen.

Ergebnisse

Es fanden sich keine Unterschiede in der Schwierigkeit der sechs Module, auch ließ sich kein Einfluss des Faktors Prüfer/in auf die Ergebnisse nachweisen. Die errechnete Reliabilität für die Prüfung in einem Modul lag bei 0,509, einem in der Literatur als „angemessen“ beschriebenen Wert.

Schlussfolgerungen

Prüfungen müssen einerseits definierten Testgütekriterien genügen, andererseits müssen sie vom personellen, organisatorischen und finanziellen Aufwand her durchführbar sein. Beide Kriterien werden von der „Freiburger Famulatureifepfung“ erfüllt.

Schlüsselwörter

kompetenzorientierte Ausbildung; kompetenzorientierte Prüfung; Performanzprüfung; Untersuchungskurs

Background

A competence-oriented and practical training of medical students requires – according to the Constructive Alignment Model – practical examinations. For the medical students of the University of Freiburg, the „clinical traineeship examination“ (Famulatureifepfung), which is a so-called performance test, completes a multistage course on basic physical examination techniques. The students have to prove their physical examination skills in one of six possible, randomly assigned examination modules. This paper investigates whether this procedure is justifiable with regard to the reliability of the examination.

Methods

After a descriptive evaluation, the examination results of the 73 students who took the clinical traineeship examination in the summer semester of 2018 were subjected to a mixed model variance analysis and a calculation of reliability.

Results

Results showed no differences in difficulty between the six modules. No influence of the factor „examiner“ on the results could be proven. The calculated reliability of the examination in one module was 0.509, a value described as „appropriate“ in the literature.

Conclusions

On the one hand, tests must satisfy defined test quality criteria, on the other hand, they must be feasible in terms of personnel, organizational and financial expenditure. Both criteria are fulfilled by the „Freiburger Famulatureifepfung“.

Keywords

competence-based training; competence-based examination; performance test; examination course

Hintergrund

Schon die Approbationsordnung von 2002 [1] hatte für Absolvent*innen des Studiengangs Medizin neben grundlegenden medizinischen, fächerübergreifenden und methodischen Kenntnissen das Beherrschen

praktischer Fertigkeiten explizit eingefordert. Dies löste seinerzeit umfangreiche Aktivitäten bei der curricularen Weiterentwicklung des Medizinstudiums in Deutschland aus.

Der im März 2017 vom Bundesgesundheitsminister, der Bundesfor-

schungsministerin sowie den Vertreter*innen der Gesundheits- und der Kultusministerkonferenz der Länder der Koalitionsfraktionen des Deutschen Bundestages beschlossene „Masterplan Medizinstudium 2020“ beschreitet diesen Weg konsequent

¹ Lehrbereich Allgemeinmedizin, Universitätsklinikum Freiburg

² Abteilung für Allgemeinmedizin, Universität Bochum

³ Medizinische Fakultät, Universitätsklinikum Heidelberg

Peer reviewed article eingereicht: 07.11.2019, akzeptiert: 20.12.2019

DOI: 10.3238/zfa.2020.0121–0126

weiter und fordert eine Neustrukturierung des Medizinstudiums in Richtung einer kompetenzorientierten und praxisnahen Ausbildung sowie praxisnaher Prüfungen [2].

Da – entsprechend dem *Constructive-Alignment-Modell* [3] – nachvollziehbar das gelernt wird, was geprüft wird, sind zur Erreichung dieser Ziele Prüfungen mit hohem praktischen Anteil ideale Steuerungselemente. Geeignete Prüfungsformate sind z.B. die *Objective Structured Clinical Examination* (OSCE) [4], die *Clinical Evaluation Exercise* (CEX) bzw. *Mini Clinical Evaluation Exercise* (Mini-CEX) [5] sowie die *Direct Observation of Procedural Skills* (DOPS) [6]. Allen diesen Formaten ist gemein, dass sie mit einem hohen zeitlichen und personellen Aufwand einhergehen und somit für ihren Einsatz bei größeren Studierenden-Kohorten problematisch sein können.

Die Medizinstudierenden der Universität Freiburg durchlaufen vom 2. Semester des 1. Studienabschnitts bis zum 1. Semester des 2. Studienabschnitts einen mehrstufigen Kurs zur Vermittlung von Basistechniken der körperlichen Untersuchung. Dieser Kurs besteht aus den fünf Modulen Kopf & Hals, Herz & Lunge, Abdomen, Wirbelsäule & Gelenke sowie Neurostatus. Die einzelnen Module werden in Kleingruppen von maximal acht Studierenden durch standardisiert ausgebildete studentische Tutoren unterrichtet. Jeweils drei Kleingruppen werden durch ärztliche Dozierende supervidiert. Der Kurs wird abgeschlossen mit einer sog. „Famulatureifeprüfung“. Diese wird in Form einer Performanzprüfung [7, 8] durchgeführt. Es wurde ein Prüfungsformat gesucht, welches ressourcensparend eine Prüfung der realen bzw. simulierten Situation zuließ, die zuvor im Untersuchungskurs unterrichtet worden war. Die Prüfung sollte den kompletten Untersuchungsablauf beinhalten und nicht nur einzelne Untersuchungsschritte wie es z.B. bei einer OSCE der Falle gewesen wäre.

Charakteristisch für diese Prüfungsform ist die von Prüfer*innen ungestörte Vorführung (= Performanz) einer praktischen Fertigkeit durch die Studierenden, hier den Untersuchungstechniken aus einem der

fünf Module. Für die Prüfung werden aus den fünf Unterrichtsmodulen sechs Prüfungsmodule: Das Modul Wirbelsäule & Gelenke wird in der Prüfung zu den beiden Modulen Wirbelsäule und Kniegelenk. Die Studierenden werden in einem jeweils zufällig zugeteilten Modul geprüft. Die Prüfung erfolgt in Freiburg als Tandemprüfung, die Studierenden wechseln sich in den Rollen Untersuchende/Untersuchte ab. Der Zeitrahmen beträgt zehn Minuten pro Studierenden, sieben Minuten für die Performanz, drei Minuten für eventuelle Nachfragen durch die Prüfer/-innen. Die Prüfenden stellen dabei Fragen zum Handlungs- und Begründungswissen. Die Fragen sind standardisiert und mit festen Punktwerten versehen.

Bewertet wird die Prüfung von einem/einer ärztlichen Prüfer*in und einem/einer Beisitzer*in aus dem Pool der studentischen *peer teacher* des Untersuchungskurses. Die Bewertung erfolgt anhand einer aus modulübergreifenden und modulspezifischen, 32 Items umfassenden Checkliste. Die modulübergreifenden Items wie Händedesinfektion oder die Inspektion haben in jedem Modul denselben maximalen Punktwert. Die modulspezifischen Items haben je nach Komplexität unterschiedliche Punktwerte, die maximale Punktzahl in jedem Modul beträgt letztlich aber wieder 32. Der Erwartungshorizont in Bezug auf die Prüfungsleistungen ist den Studierenden über schriftlich ausformulierte Standards bekannt, die Checklisten werden nicht transparent gemacht. Die Bestehensgrenze liegt bei 60 %.

Ziel der hier vorgestellten Studie war es u.a. zu überprüfen, ob die unterschiedlichen Module unterschiedliche Schwierigkeitsgrade aufweisen und ob sich der zweite Prüfungsdurchgang vom ersten hinsichtlich der Höhe der erreichten Punktzahl unterscheidet. Zudem galt es zu zeigen, dass der Faktor „Prüfer*in“ keinen signifikanten Einfluss auf das Prüfungsergebnis hat.

Des Weiteren sollte untersucht werden, ob die Prüfung in nur einem Modul ausreichend ist, um eine reliable Aussage über die Fertigkeiten der Studierenden im Hinblick auf die körperliche Untersuchung treffen zu können.

Methoden

Stichprobe

Die Erhebung der Daten erfolgte im Rahmen der im Sommersemester 2018 durchgeführten „Famulatureifeprüfung“. Abweichend von der regulären Prüfung wurden die Studierenden im Rahmen der Studie nicht in einem, sondern in zwei zufällig ausgewählten Modulen geprüft. Diese Vorgehensweise wurde gewählt, um die Reliabilität der Prüfung berechnen zu können. Damit verdoppelte sich die Prüfungszeit für die sechs eingesetzten Prüfungsteams (s.u.) von ca. drei auf ca. sechs Stunden. Die Daten aller 73 Studierenden, die an dieser Prüfung teilgenommen hatten, gingen in die Auswertung ein.

Erhebungsinstrument

Die Umsetzung und Auswertung der Prüfungen erfolgte mit digitaler Unterstützung. Mithilfe von I-Pads und der App tOSCE konnten die Prüfungsschecklisten elektronisch ausgefüllt werden, die Daten standen nach Abschluss der Prüfung automatisch in Form von Excel-Tabellen zur Verfügung.

Datenverarbeitung und Datenanalyse

In einem ersten Schritt wurden die Prüfungsdaten pseudonymisiert. Die Daten wurden mithilfe der Statistikprogramme IBM SPSS Statistics (Version 25) und R Version 3.5.1 ausgewertet. Zunächst erfolgte eine deskriptive Datenauswertung durch Darstellung von Mittelwert, Standardfehler, Standardabweichung, Minimum und Maximum.

Die Beantwortung der Frage nach der Vergleichbarkeit der Module sowie des Abschneidens der Studierenden in der ersten bzw. zweiten Prüfung der Freiburger „Famulatureifeprüfung“ erfolgte mithilfe einer Mixed-Model-Varianzanalyse. Feste Faktoren waren „Modul“ und „Prüfungsart“ (Erst-/Zweitprüfung). Zufällige Faktoren waren „Studierende*r“ und „Prüfer*in“.

Aus den resultierenden Varianzkomponenten lässt sich die Reliabilität einer Prüfung bei n festen Stationen und wechselnden Prüfern mittels der Formel

Modul	Erstprüfung			Zweitprüfung			Erst- und Zweitprüfung		
	x in Punkten	SEM	SD	x in Punkten	SEM	SD	x in Punkten	SEM	SD
Abdomen (A)	27,30 (85,31 %)	0,86	3,37	26,40 (82,50 %)	1,15	3,63	26,94 (84,19 %)	0,69	3,43
Herz & Lunge (HL)	26,41 (82,53 %)	1,05	3,65	26,81 (83,78 %)	1,21	4,36	26,62 (83,19 %)	0,79	3,95
Kopf & Hals (KH)	26,25 (82,03 %)	1,15	4,14	28,24 (88,25 %)	0,83	2,75	27,16 (84,88 %)	0,74	3,64
Knie (K)	26,38 (82,44 %)	1,13	3,92	26,29 (82,16 %)	1,26	4,38	26,33 (82,28 %)	0,83	4,06
Neurostatus (N)	23,10 (72,19 %)	1,32	4,18	26,70 (83,44 %)	0,75	2,81	25,20 (78,75 %)	0,78	3,82
Wirbelsäule (WS)	25,91 (80,97 %)	1,27	4,24	26,35 (82,34 %)	0,87	3,15	26,15 (81,72 %)	0,74	3,61
Summe	32 (100 %)			32 (100 %)			32 (100 %)		

x: Mittelwert, SEM: Standardfehler des Mittelwerts, SD: Standardabweichung

Tabelle 1 Darstellung der modulspezifischen Ergebnisse

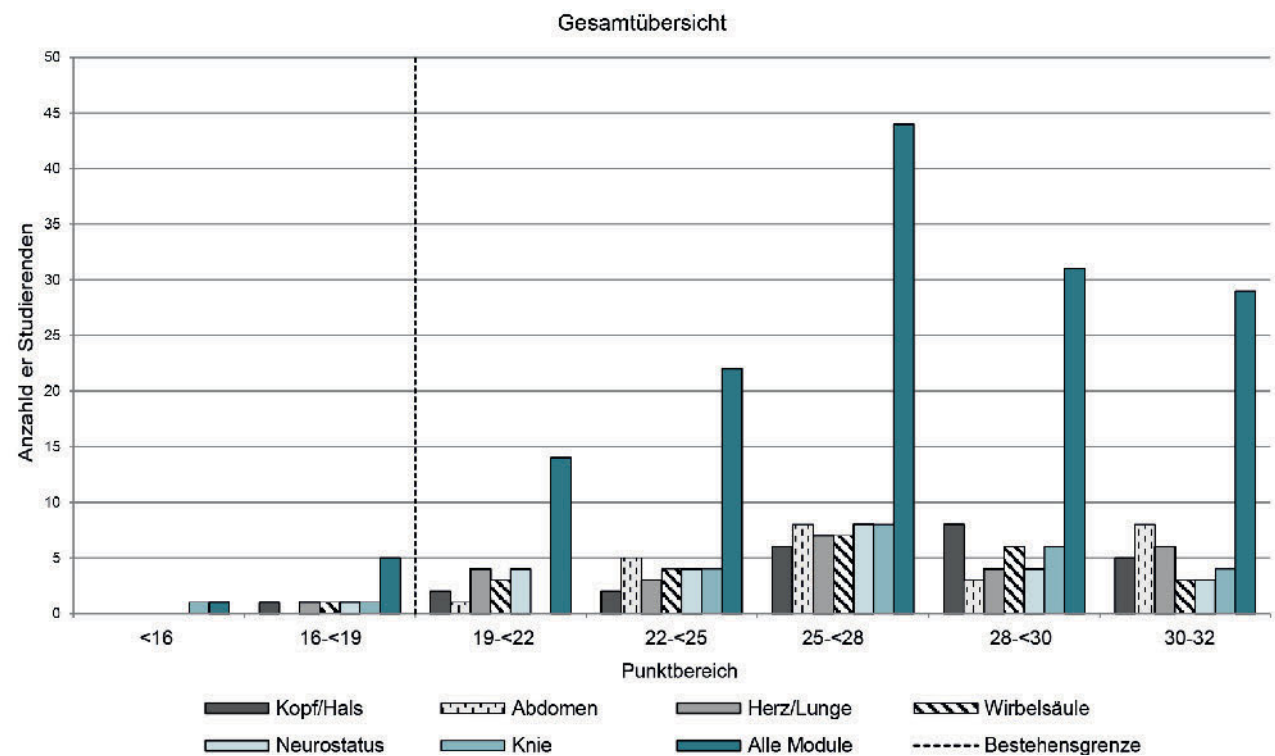


Abbildung 1 Übersicht über die Punkteverteilungen der einzelnen Prüfungsmodule

$$\frac{Var(Stud)}{Var(Stud) + Var(Prüfer)/n + Var(Residual)/n}$$

bestimmen (vgl. hierzu die Definition des Generalisierbarkeitskoeffizienten in der Generalisierbarkeitstheorie [9]).

Ergebnisse

Beschreibung der Stichprobe

Die Gesamtstichprobe umfasste 73 Studierende, 48 weiblich (65,75 %) und 25 männlich (34,25 %). Der Al-

tersdurchschnitt betrug 25,18 Jahre (SD 3,99). Die jüngsten Studierenden waren 20 Jahre alt, der/die älteste Student/in war 40 Jahre alt. Im Durchschnitt waren die männlichen Studierenden etwa ein Jahr älter (x = 25,88 Jahre (SD 4,18)) als ihre weiblichen Kommilitoninnen (x = 24,81 Jahre (SD 3,89)).

Beschreibung der Ergebnisse

Es wurden bei 73 teilnehmenden Studierenden insgesamt 146 Prüfungen durchgeführt. 93,15 % der Studieren-

den haben die Prüfung bestanden. Die Verteilung der sechs Module erfolgte in der Gesamtheit und im Hinblick auf Erst- bzw. Zweitprüfung nahezu gleichmäßig. Auch die 15 möglichen Modulkombinationen waren gleich verteilt.

Die Studierenden erreichten in der Gesamtprüfung im Mittel eine Punktzahl von x = 52,81 Punkten, dies entspricht 82,52 % der maximalen Punktzahl. Betrachtet man die Ergebnisse jeder Einzelprüfung, wurde hier ein Mittelwert von x = 26,41

(82,53 %) Punkten von maximal 32 Punkten erreicht. In der Erst- und Zweitprüfung konnte jeweils eine maximale Punktzahl von 32 Punkten erreicht werden. Der Mittelwert in der Erstprüfung lag bei $x = 26,03$ Punkten (81,34 %) und in der Zweitprüfung bei $x = 26,78$ Punkten (83,69 %). Die modulspezifischen Ergebnisse sind Tabelle 1 zu entnehmen.

Die Häufigkeitsverteilungen der erreichten Punktzahlen innerhalb der einzelnen Module sind in Abbildung 1 dargestellt. Werden alle 146 Prüfungen modulübergreifend zusammengefasst, zeigt sich ein nahezu kontinuierlicher Anstieg der Studierendenzahl ausgehend vom Punktbereich < 16 hin zum Punktbereich 25–28. Danach nahm die Studierendenzahl hin zum Punktbereich 30–32 wieder ab.

Das Box-Plot-Diagramm (Abb. 2) lässt Aussagen über Verteilung, Streuung, die Form der Verteilung und über Ausreißer zu. Der Median der jeweiligen Module befindet sich im Bereich zwischen 25 und 28 Punkten. In fünf der sechs Module lagen die mittleren 50 % der Daten im Punktbereich zwischen 24 und 30 Punkten. Einzig im Modul Neurostatus lagen die mittleren 50 % der Daten bereits bei 22 bis 28 Punkten. Unterschiede zeigten sich in der Spannweite der erhobenen Daten je Modul. Die Module Abdomen, Kopf & Hals, Neurostatus und Wirbelsäule wiesen eine Spannweite von ca. 12–14 Punkten auf. Die größte Spannweite fand sich im Modul Herz & Lunge mit ca. 15 Punkten. Die geringste Spannweite mit ca. 8,5 Punkten wies das Modul Knie auf, wobei hier zwei Ausreißer bei 15,5 und 16 Punkten auftraten.

Korrigiert man die Daten um den Faktor Prüfer/innen (s. Methodik), zeigen sich im Vergleich zu den Verteilungen in Abbildung 2 keine signifikanten Abweichungen.

Vergleich der verschiedenen Prüfungsmodule

Mithilfe der durchgeführten Varianzanalyse nach dem gemischten Modell (*mixed model*) wurden die erhobenen Daten auf signifikante Unterschiede zwischen den sechs Modulen bezüglich ihrer Schwierigkeit bzw. die Erst-

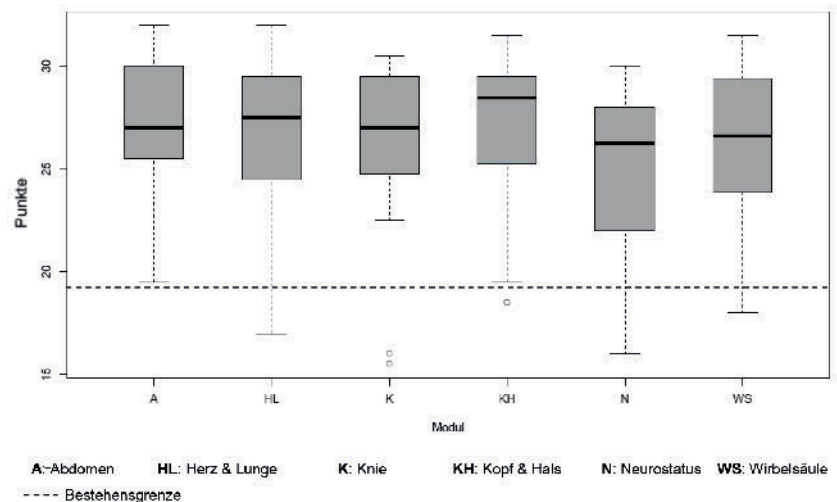


Abbildung 2 Punkteverteilung modulspezifisch

Anzahl der Module	Reliabilität
1 Modul	0,509
2 Module	0,674
3 Module	0,756*
4 Module	0,805*
5 Module	0,838*
6 Module	0,861*

* Bei den mit Sternchen gekennzeichneten Ergebnissen handelt es sich um hypothetische Berechnungen.

Tabelle 2 Ergebnisse der Reliabilitätsberechnung

und Zweitprüfung auf signifikante Unterschiede überprüft.

Dabei fanden sich keine signifikanten Unterschiede in der Schwierigkeit der sechs Module ($p = 0,461$), es konnte aber gezeigt werden, dass das Ergebnis der Zweitprüfung um statistisch signifikante 0,75 Punkte höher lag als bei der Erstprüfung ($p = 0,046$).

Der Faktor Prüfer/in erwies sich als nicht signifikant ($p = 0,061$), es ließ sich also kein verzerrender Einfluss der Prüfer/innen auf die Prüfungsergebnisse zeigen.

Reliabilität der „Famulatureifepfung“

Die Berechnung der Reliabilität erfolgte nicht nur für die im Sommersemester 2018 umgesetzte Prüfung

von zwei zufällig gewählten Modulen, sondern auch für die Varianten, wenn die Studierenden nur in einem bzw. hypothetisch auch in drei bis sechs zufällig gewählten Prüfungen geprüft worden wären. Die Tabelle 2 zeigt die Ergebnisse dieser Reliabilitätsberechnungen. Die errechnete Reliabilität der Prüfung eines Moduls liegt bei 0,509, für die Prüfung von zwei Modulen bei 0,674. Mit steigender Anzahl der geprüften Module steigt auch die errechnete Reliabilität. Würden alle sechs möglichen Module geprüft werden, würde die Reliabilität einen Wert von 0,861 annehmen.

Diskussion

Ziel dieser Studie war es, zu untersuchen, ob die „Freiburger Famulatureifepfung“ in ihrem Aufbau für

eine ausreichend faire Beurteilung der Basis-Untersuchungstechniken geeignet ist und somit an diesem Prüfungsformat weiterhin festgehalten werden kann.

Es konnte gezeigt werden, dass sich die sechs Module hinsichtlich der erzielten Prüfungsergebnisse statistisch nicht signifikant unterschieden. Ein geringer, aber signifikanter Unterschied fand sich bei den Ergebnissen zwischen dem 1. und 2. Prüfungsmodul: Die Zweitprüfung fiel um durchschnittlich 0,75 Punkte besser aus.

Die errechnete Reliabilität bei Prüfung in nur einem Modul lag bei 0,509.

Vergleich der Module

Um zu validen Ergebnissen zu kommen, erfolgte die Zuordnung der sechs Module auf die 146 Prüfungen mit dem Ziel, diese in ihrer Gesamtanzahl und mit ihren Kombinationsmöglichkeiten möglichst gleichmäßig zu verteilen.

Als Voraussetzung der Repräsentativität der Prüfungsleistung in einem der sechs Module im Hinblick auf die Beurteilung der Fähigkeit zur körperlichen Untersuchung insgesamt muss ein vergleichbarer Schwierigkeitsgrad aller Prüfungsmodule gelten. Eine Herausforderung lag in der Ausarbeitung der Aufgabenstellungen der einzelnen Module. Der vorgegebene zeitliche Rahmen von zehn Minuten ermöglichte einen strukturierten Ablauf, führte aber auch dazu, dass einige Module (z.B. Kopf & Hals) in ihren Prüfungsinhalten reduziert werden mussten. Das Modul Kopf & Hals wies schließlich die im Durchschnitt höchste erreichte Punktzahl auf. Ein Grund dafür kann gewesen sein, dass die Aufgabenstellung in diesem Modul durch die dezidiertere Formulierung der einzelnen Untersuchungsschritte durch die Studierenden leichter abzuarbeiten war. Die selbstständige Reproduktion und Strukturierung des Untersuchungsablaufs war in anderen Modulen vergleichsweise anspruchsvoller. Letztlich aber deutet die Tatsache, dass die Ergebnisse keinen statistisch signifikanten Unterschied zwischen den einzelnen Modulen aufwiesen, auf die Ausgewogenheit der Module

hin. Dies ist die wesentliche Bedingung, das Format der „Freiburger Famulaturreifepfung“ in der aktuellen Form beizubehalten.

Die Ergebnisse der Prüfung im jeweils zweiten Modul fiel um signifikante 0,75 Punkte höher aus, ohne dass sich die Ausgewogenheit in der Schwierigkeit der Module untereinander geändert hätte. Dies mag als „Lerneffekt“, als Gewöhnung an das Prüfungsformat gedeutet werden.

Reliabilität der Prüfung

Die Reliabilität erreichte einen Wert von 0,509 bei der Prüfung eines Moduls. Dies entspricht nach Cicchetti [10] einem „angemessenen“ Niveau an klinischer Signifikanz. Werte zwischen 0,60 und 0,74 wären gut, zwi-



Lena Krabbe ...

.... studierte Medizin an der Albert-Ludwigs-Universität Freiburg. Seit 2017 ist sie wissenschaftliche Mitarbeiterin im Lehrbereich Allgemeinmedizin des Universitätsklinikums Freiburg. Im Rahmen ihrer Promotion im Lehrbereich beschäftigte sie sich mit dem Aufbau und der Umsetzung von praktischen Prüfungen und deren Stellenwert im Medizinstudium.

schen 0,75 und 1,00 exzellent. Unter Berücksichtigung der Tatsache, dass eine Überprüfung der Kompetenz von Studierenden per se einen hohen Einsatz an Ressourcen erfordert und andererseits die personellen wie finanziellen Mittel des Lehrbereichs Allgemeinmedizin der Uniklinik Freiburg begrenzt sind, stellt die „Freiburger Famulaturreifepfung“ einen pragmatischen Ansatz zur adäquaten Überprüfung der Basis-Untersuchungstechniken dar. Die Prüfung in zwei Modulen, wie im Rahmen dieser Studie durchgeführt, würde die Reliabilität auf ein „gutes“ Signifikanz-

niveau heben, wäre aber für die gesamte Studierendenkohorte eines Jahrganges mit einem nicht zur Verfügung stehenden personellen und organisatorischen Aufwand verbunden.

Schlussfolgerungen

Die Fähigkeit zur körperlichen Untersuchung steigt nicht automatisch mit dem Voranschreiten der universitären Ausbildung. Auch Praktika garantieren keinen zwangsläufigen Zuwachs an Kompetenzen zur körperlichen Untersuchung. So erscheint es wichtig, eine curricular gesteuerte Lehre von praktischen Fertigkeiten zu etablieren und diese auch zu überprüfen [11, 12]. Um den Studierenden gerecht zu werden, sollte eine solche Prüfung definierte Gütekriterien erfüllen. Andererseits muss sie auch für die Lehrenden ressourcenschonend umsetzbar sein. Beide Kriterien werden von der „Freiburger Famulaturreifepfung“ erfüllt.

Interessenkonflikte:

Keine angegeben.

Literatur

1. Approbationsordnung für Ärzte vom 28. Oktober 1970 (SGBI 1 5.1458). https://download.uni-mainz.de/verwaltung-sl/ordnungen/AppO_Medizin_bis_09_03.pdf (letzter Zugriff am 19.07.2019)
2. Gemeinsame Pressemitteilung, Wichtiger Schritt zu modernem Medizinstudium: „Masterplan Medizinstudium 2020“. www.allgemeinmedizin.uni-frankfurt.de/aktuelles/2017/170331_PM_Masterplan_Medizinstudium.pdf (letzter Zugriff am 02.03.2020)
3. Biggs J. Enhancing teaching through constructive alignment. *High Educ* 1996; 32: 347–364
4. Harden R, Stevenson M, Wilson DW, Wilson GM. Assessment of clinical competence using objective structured examination. *BMJ* 1975; 1: 447–451
5. Norcini JJ, Blank L, Arnold G, Kimball H. The mini-cex (clinical evaluation exercise): a preliminary investigation. *Ann Intern Med* 1995; 10: 795–799
6. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE guide no. 31. *Med Teach* 2007 29: 9–10, 855–871
7. Gage N, Berliner D. Lehrerbeurteilung und Noten. In: Beltz (Hrsg.). Pä-

- dagogische Psychologie, 5. Aufl. Weinheim: Psychologie Verlags Union, 1996: 635–639
8. Klemme B, Andres J, Geuter G, et al. Performanzprüfung. In: Klemme B (Hrsg.). Lehren und Lernen in der Physiotherapie, 1. Aufl. Stuttgart: Thieme, 2012: 158
9. Brennan RL. Generalizability theory. New York: Springer, 2001
10. Cicchetti DV. Guidelines, criteria and rules of thumb for evaluating normed and standardized assessment instrument in psychology. Psychol Assess 1994; 6: 284–290
11. Krautter M, Diefenbacher K, Koehl-Hackert N, et al. Short communication: final year students' deficits in physical examination skills performance in Germany. Z Evid Fortbild Qual Gesundhwes 2015; 109: 59–61
12. York NL, Niehaus AH, Markwell SJ, Folse JR. Evaluation of students' physical examination skills during their surgery clerkship. Am J Surg 1999; 177: 240–243

Korrespondenzadresse

Dr. med. Klaus Böhme
Abteilung für Allgemeinmedizin
Ruhr-Universität Bochum
Universitätsstr. 150,
44801 Bochum
Tel.: 0234 32–24894
klaus.boehme@rub.de



45. wissenschaftliches GHA-Symposium

„Zukünftige Handlungsfelder der Aus- und Weiterbildung in der Allgemeinmedizin“

Termin: 16./17. Mai 2020

Veranstaltungsort: Stiftung RUFZEICHEN GESUNDHEIT
Wolfratshauer Straße 9, 82065 Baierbrunn bei München
Programm und Hotelbuchung unter www.gha-info.de

Wir freuen uns auf Ihre Teilnahme, interessante Beiträge und engagierte Diskussionen im anregenden Baierbrunner Ambiente.