Aus dem Lernzentrum der Charité
und der Klinik für Gynäkologie und Geburtshilfe
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Verbesserung der medizinischen Lehre durch Einbezug zukunftsrelevanter Praxisinhalte und nicht-kognitiver Faktoren in den medizinischen Lehr-Lernprozess

zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Therese Schröder
aus Marburg

Datum der Promotion: 07.12.2018

Inhaltsverzeichnis

Abkürzungsverzeichnis

CLT	Cognitive Load Theory
DACH-Region	Deutschland, Österreich, Schweiz
ESS	Experiential Shame Scale
GI-Trakt	Gastrointestinal-Trakt
GRS	Global Rating Scale
NKLM	Nationaler Kompetenzbasierter Lernzielkatalog Medizin
OF	Outcome Framework
OSCE	Objectively Structured Clinical Examination
SP	Simulationspatient/Simulationspatientin
TOSCA-3	Test of Self Conscious Affect, Version 3

Abstrakt in Deutsch

In einer ersten Studie befasse ich mich mit der Lehre praktischer Fertigkeiten während der medizinischen Ausbildung, um eine Aussage unterstützen zu können, welche praktischen Fertigkeiten in Zukunft von Relevanz sein werden. In einer zweiten Studie beschäftige ich mich mit der Selbstsicherheit, welche durch Simulationstrainings signifikant erhöht werden kann. Die dritte Studie bezieht sich auf das Erlernen praktischer Fertigkeiten am Beispiel der Brustuntersuchung der Frau und auf den Einfluss von Schamgefühlen auf die Performance.

Der ersten Studie liegt eine Expertenbefragung zugrunde, die zur Entwicklung von elf Thesen zum Gesundheitswesen 2025 führte. In einer sich anschließenden, zweistufigen explorativen Delphi-Methode wurden 288 Lernziele im Hinblick auf ihre zukünftige Relevanz für die ärztliche Ausbildung einem Bewertungsprozess unterzogen. Insgesamt wurden 231 Lernziele als relevant und 57 Lernziele als irrelevant bewertet. Die Delphi-Methode mit ihren Ergebnissen soll dazu dienen, auch zukünftige Curricula in Einklang mit den sich immer verändernden Anforderungen an die Absolventen des Medizinstudiums zu bringen.

Die zweite Studie basiert auf einer Nachtdienst-Simulation mit sieben Stationen und 30 Teilnehmenden. Die Fünferteams bestanden aus Teamleiter, Teammitgliedern und Beobachtern. Das Gefühl des Vorbereitetseins wurde zu Beginn und nach fünf Tagen per Fragebogen erhoben; „Selbstsicherheit“ wurde nach jeder durchlaufenen Station erhoben. Für das Einschätzen der Selbstsicherheit war die Rolle innerhalb der Gruppe unerheblich, insgesamt hat die Simulation die Selbstsicherheit verbessert.

Bei der dritten Studie handelt es sich um eine randomisierte, kontrollierte Studie mit 49 Medizinstudierenden, die alle eine theoretische Einführung zur Untersuchung der weiblichen Brust erhielten. Danach erfolgte die Aufteilung in zwei Gruppen (Gruppe A übte mit Puppe, Gruppe B mit Simulationspatientin). Abschließend fand eine einheitliche Prüfung mit einer Simulationspatientin mit pathologischem Brustmodell statt. Variablen waren das situative Schamgefühl und die Untersuchungsqualität. Es lassen sich Schamgefühle hervorrufen und diese durch geeignete Trainingsvarianten im Verlauf verändern. Statistisch signifikant korreliert in einer Teilgruppe die Abnahme des situativen Schamgefühls mit dem besseren Auffinden pathologischer Befunde. Zukünftige Simulationen sollten die emotional affektiven Komponenten des praktischen Handelns beachten, um sie für den Lernprozess fruchtbar machen zu können.

Abstrakt in Englisch

My first study deals with practical skills in undergraduate medical education in order to make a prediction of which practical skills will be relevant in the future. A second study addresses students' self-efficacy and how it can be significantly improved by simulation trainings. The third study deals with the acquisition of intimate examination skills for the clinical breast exam and the effect of shame on subsequent performance.

The first study is based on consultation with experts leading to 11 theses about general, future developments in healthcare by 2025. This was followed by a two-stage, exploratory Delphi-process to evaluate 288 learning goals and their future relevance for medical education. A total of 231 learning goals were deemed relevant; 57 learning goals were rated as irrelevant. The Delphi-method and the results provide a guide for future curricula and the ever-changing requirements for medical school graduates.

The second study is based on a night shift simulation with seven cases and 30 participants. The groups of five students consisted of a leader, members and an observer. Feelings of preparedness were measured with a questionnaire at baseline and five days after the event; confidence was surveyed after every completed case. The role within the group was negligible for estimating confidence. Overall self-efficacy improved with the simulation.

The third study is a randomised controlled study with 49 medical students who all received basic clinical breast examination training. Afterwards they were split into two groups (group A practiced with a mannequin, group B with a standardized patient). Finally, all students went through the same exam with a standardized patient wearing a pathological strap-on breast model. Variables were measures of situational shame and outcome and process quality. It was possible to elicit feelings of shame and change them over the course of training. A decline of state shame correlated significantly with a better performance in finding breast lumps. Future simulations should include affective components relevant to training in order to improve learning outcomes.

a. Einführung

Die medizinische Lehre unterliegt einem ständigen Wandel und muss sich laufend anpassen: an veränderte Bedürfnisse der Studierenden[A], an neue Technologien, an Gelder & Personal, an Leistungsaufträge durch die Universitäten und Städte, an neue wissenschaftliche Erkenntnisse und vieles mehr. Daraus ergibt sich für die Forschung im Bereich der medizinischen Lehre ein großes Feld an offenen Fragen. Zudem zeigt sich eine Diskrepanz zwischen Anforderungen an den approbierten Arzt und die aktuellen Curricula des Medizinstudiums, deren Überarbeitung und Implementation oft Jahre dauert. Obwohl inzwischen die Vermittlung praktischer Fertigkeiten in der DACH-Region Einzug in das Medizinstudium gehalten hat, fühlen sich viele Absolventen nicht ausreichend auf den Arztberuf vorbereitet [1]. OFs stellen den Versuch dar, die Anforderungen und zu erwerbenden Fertigkeiten festzuhalten; diese weisen jedoch häufig große Unterschiede in Inhalt und Struktur auf [2, 3].

Es gibt viele Berufsanfänger, die nur über eine ungenügende Selbstsicherheit in Bezug auf ihre Diagnosestellung verfügen – unabhängig davon, wie gut/exakt die Diagnose gestellt wurde [4]. Ein bestimmtes Maß an Selbstsicherheit ist von Bedeutung für die Weiterbehandlung von Patienten, wie schnell etwa die nächsten Maßnahmen, zusätzliche Diagnostik oder Therapien eingeleitet werden. Situationsabhängig ausreichende Selbstsicherheit ist demnach ein entscheidender Punkt dafür, dass wichtige Maßnahmen auch wirklich umgesetzt werden [5, 6]. Die Einschätzung, sich ausreichend vorbereitet zu fühlen, hängt von vielen verschiedenen Faktoren ab: unter anderem vom Ausbildungsort, dem Einbezug von Feedback, der Mischung von Theorie und Praxis verbunden mit ausreichend Zeit zum Üben und Diagnosestellen. In der Simulation sollten diese Faktoren alle integriert werden und somit nicht nur das individuelle Gefühl der Selbstsicherheit erhöhen, sondern zudem zur Steigerung der Objective Performance führen [7]. Zudem ist erwiesen, dass eine Beobachter-Rolle der Studierenden in der Simulation vergleichbare praktische Lerneffekte hat wie die aktive Rollenübernahme [8].

Es gibt viele Studien, die sich mit dem Erlernen motorischer Fertigkeiten generell und mit praktischen Fertigkeiten in der medizinischen Ausbildung im Speziellen beschäftigen. Im Gegensatz zu hauptsächlich manuellen Tätigkeiten wie Wundversorgung oder die venöse Blutentnahme, gibt es in der Medizin auch viele praktische Fertigkeiten, die verschiedene Grade an Selbstreflexion erfordern und – je nach Aufgabe – eine veränderte Selbstwahrnehmung

[A] Um die Lesbarkeit dieser Arbeit zu verbessern, werden Personengruppen in einer Gender-übergreifenden Form bezeichnet (Studierende, Ärzte, Patienten). Dabei sind, wenn nicht explizit anders angegeben, stets alle Geschlechter gemeint.

auslösen können. Diese praktischen Fertigkeiten, wie beispielsweise die gynäkologische Untersuchung oder die digital-rektale Untersuchung, können eine ganze Reihe von Emotionen auslösen [9], sowohl beim Untersucher als auch beim Patienten. Während diese selbstreflexiven Prozesse im Bereich der Psychologie schon länger bekannt sind, gibt es in der medizinischen Lehre noch viele offene Fragen bezüglich der Auswirkungen eben dieser Emotionen und selbstreflexiven Prozesse auf die praktischen Fertigkeiten der Medizinstudierenden und die potentiell weitreichenden Folgen auf das nachfolgende Arbeitsleben eines Arztes [10] und somit auch der Patientenversorgung. Im Kontext der medizinischen Lehre sind die Auswirkungen von Emotionen und selbstreflexiven Prozessen auf das Erlernen praktischer Fertigkeiten bisher wenig erforscht [11, 12]. Die Hauptemotion, mit der wir uns in unserer dritten Studie befassen, ist das Schamgefühl. Dieses muss man von Peinlichkeit und Schuldgefühlen unterscheiden, auch wenn diese drei Emotionen umgangssprachlich häufig synonym verwendet werden: Eine Situation kann einem peinlich erscheinen, wenn oberflächliche soziale Normen missachtet werden, häufig kann man über dieses Erlebnis später jedoch lachen und die negativen Emotionen haben sich bereits verflüchtigt. Schuld und Schamgefühl hingegen sind sogenannte selbstreflexive Emotionen, die entstehen, wenn eigene innerste Werte missachtet werden – schuldig kann man sich für ein bestimmtes Verhalten fühlen, während man sich schämt für sein Innerstes/die eigene Person. Schamgefühle können unterschieden werden: Zum einen geht es um ein grundlegendes Persönlichkeitsmerkmal, zum anderen um ein situatives Gefühl. In einer gegebenen Situation schäm sich eine Person mit einem erhöhten Niveau an Schamgefühl als Persönlichkeitsmerkmal eher als jemand mit einem niedrigen Niveau – dieser fühlt sich wahrscheinlich eher schuldig. Es bestehen große individuelle Unterschiede; diese sind wichtig, wenn man sich die daraus resultierenden Verhaltensmuster anschaut: Ausgeprägte Schamgefühle fördern einerseits ein starkes Vermeidungsverhalten [13], welches in der medizinischen Lehre kontraproduktiv erscheint, können jedoch andererseits zu einer aktiveren Verarbeitung der Situation führen und somit den Lerneffekt im Sinne eines Erinnerungseffektes steigern [10].

Aus Sicht der CLT können Gefühle für den Lernprozess hinderlich oder förderlich sein [14]. Das hängt direkt von der Strukturierung dieses Prozesses ab: Zuerst geht es um einen angemessenen Umgang mit der Komplexität der Aufgabenstellung (intrinsic load; dabei können unterkomplexe Teillernschritte ebenso problematisch sein wie zu komplexe Teillernschritte – in beiden Fällen ist die sichere Orientierung am Lernziel für die Lernenden nicht gegeben), zum Zweiten ist für den Erfolg die Wahl der Lehr-/Lernmethoden von großer Bedeutung (extraneous load; in welchem Rahmen die Aufgabe zu bewältigen ist), und zum Dritten geht es um die nachhaltige Wirksamkeit des Gelernten im Sinne der Abrufbarkeit des erworbenen Wissens und der

Handlungsmuster (germane load; aktive Verfügbarkeit von Gelerntem im Langzeitgedächtnis). Für den Umgang mit Emotionen bedeutet dies, dass sie in Qualität und Quantität als Bestandteil des Lernprozesses ausdrücklich zu berücksichtigen sind, damit sie ihre konstruktive Wirkung für das Lernen entfalten können.

Zielstellung

Im Rahmen dieser Arbeit habe ich mich mit Fragestellungen aus dem Bereich der medizinischen Lehre, deren Anforderungen in der Zukunft, der simulationsbasierten Lehre, Selbst-Assessment und Selbst-Effizienz sowie der Vermittlung praktischer Fertigkeiten und Lehrinhalten anhand einer gynäkologischen Vorsorgeuntersuchung der Brust und deren Kopplung an Emotionen beschäftigt. Die Ziele meiner Arbeit waren:

- Hypothesen zu Trends im Gesundheitssektor auf die Weise zu entwickeln, dass sie angenommene Anforderungen abbilden, dass sie einer Überprüfung unterzogen werden und dass mit Hilfe einer empirisch validierten Bewertung zukunftsfähige Ausbildungsinhalte identifiziert werden konnten.
- In einem für den beruflichen Alltag bedeutsamen Ausschnitt, dem Nachtdienst, eine Simulation anzubieten, um zum einen die Selbstsicherheit und zum zweiten das Gefühl des Vorbereitetseins zu befördern, auch unter Berücksichtigung der spezifischen, in der Simulation von jedem Einzelnen übernommenen Rolle.
- Es sollte das situative Schamgefühl hervorgerufen, anhand einer Brustuntersuchung der Frau mit der kongruenten Emotion des Schamgefühls des Untersuchenden gekoppelt und der Effekt für die praktischen Fertigkeiten registriert werden. Als theoretischer Rahmen diente die CLT.

b. Methodik

In der ersten Studie wurden mittels teilstrukturierter Interviews von medizinischen Experten Thesen bezüglich der allgemeinen zukünftigen Entwicklung in der Medizin entwickelt und diese in einer ersten Vorbereitungsphase hinsichtlich ihres wahrscheinlichen Eintretens bis zum Jahr 2025 gewichtet. In einer zweiten Arbeitsphase wurden die Lernziele des deutschen OF, auch bekannt als NKLM, durch Ärzte aus verschiedensten Einrichtungen und Fachrichtungen in ganz Deutschland hinsichtlich ihrer mittelfristigen Relevanz im Medizinstudium in zwei Durchgängen mit Hilfe der Delphi-Methode bewertet. Dazu bewerteten alle Teilnehmer die Eintrittswahrscheinlichkeit der in der ersten Phase generierten Thesen, danach wurden zufällige Gruppen eingeteilt, welche jeweils einen Teilbereich mit ca. 30 Lernzielen aus dem Konsensusstatement bewerteten. Nach einer Zwischenauswertung durch die Forschungsgruppe wurde im nächsten Schritt eine zufällige Einteilung der Teilnehmer in zwei Gruppen vorgenommen und die noch verbleibenden Lernziele, ca. 50/Gruppe, erneut bewertet. Die Details zur Methodik sind in der ausgewählten Publikation näher beschrieben [15].

In einer weiteren Studie nahmen 30 Studierende im letzten Studienjahr der Charité an einem Nachtdienst auf einer simulierten Rettungsstelle teil. Alle Teilnehmer gaben ihr schriftliches Einverständnis zur Auswertung und Publikation der Ergebnisse. Der Datenschutzbeauftragte der Charité war über die Datenerhebung informiert. Ein Ethikvotum war nicht notwendig, da keine Patienten involviert waren. In randomisierten Fünfer-Gruppen durchliefen die Studierenden sieben verschiedene Fallstationen und wechselten dabei die Rollen zwischen Teamleiter, Teamteilnehmer und Beobachter. Es waren außerdem Tutoren anwesend, die den jeweiligen Fall betreuten sowie den Ablauf unterstützten. Das Gefühl, in verschiedenen Fachdisziplinen der Medizin vorbereitet zu sein, wurde zu Beginn des Nachtdienstes und nach fünf Tagen mit Hilfe eines Fragebogens gemessen. Nach jeder Fallstation wurde zudem die individuelle Selbstsicherheit angegeben, das Team schätzte die Selbstsicherheit des Teamleiters ein, es wurde ein Feedback in der Gruppe gegeben und der Fall wurde evaluiert. Eine tiefergehende Darstellung der Methoden findet sich in der ausgewählten Publikation [16].

In einer weiteren Studie wurden 49 Studenten aus dem dritten und vierten Studienjahr der Charité nach einer gemeinsamen theoretischen Einführung durch Tutoren des Lernzentrums zum Thema klinische Brustuntersuchung der Frau in zwei Gruppen aufgeteilt. Alle Teilnehmer gaben ihr schriftliches Einverständnis zur Auswertung und Publikation der Ergebnisse. Ein gültiges Ethikvotum lag vor. Teilnehmer der Gruppe A trainierten in Kleingruppen an einer Puppe mit

umgeschnalltem Brustmodell ohne Pathologien, Teilnehmer der Gruppe B trainierten in Kleingruppen an den echten Brüsten einer gesunden Simulationspatientin. In beiden Gruppen betrug die Übungszeit 90 Minuten, es wurde Feedback gegeben, die Kleingruppen bestanden aus drei bzw. vier Studierenden, unterrichtet wurde durch einen studentischen Tutor. In der Gruppe ohne Simulationspatientin wurden zudem noch zwei Texte über emotionale Aspekte der Arzt-Patienten-Beziehung gelesen und eine kurze Diskussion über Schamgefühle geführt. Zwei Tage danach durchliefen beide Gruppen dieselbe Prüfung in einer simulierten Arztpraxis mit Simulationspatientin und umgeschnalltem Brustmodell mit Pathologien. Dabei wurden die Studierenden gefilmt; die Auswertung wurde durch vier Fachärzte vorgenommen, welche über Details der Studie nicht informiert waren und nicht wussten, welcher Interventionsgruppe die Studierenden angehörten. Dabei kamen zwei unterschiedliche Bewertungsskalen zum Einsatz: die kanadische GRS und die OSCE Checkliste [17, 18, 19]. Zu mehreren Zeitpunkten wurden zudem verschiedene Fragebögen eingesetzt, um das generelle Niveau von Schamgefühlen und das situative Schamgefühl der Studierenden sowie mögliche Confounder zu messen. Die Details zu den Methoden finden sich in der ausgewählten Publikation [20].

c. Ergebnisse

In unserer ersten Studie konnten wir zeigen, dass viele der auch in Zukunft wichtigen ärztlichen Fertigkeiten, beispielsweise für die Behandlung Alters-assoziierter Krankheiten wie Demenz, ein großes Maß an Soft Skills und kommunikativen Fähigkeiten von Seiten der behandelnden Ärzte benötigen. Ebenso spielt das interprofessionelle Arbeiten eine große Rolle und sollte besser in die medizinische Lehre integriert werden.

a. Teilnehmer: 651 Experten bewerteten die Zukunftsthesen im ersten Durchgang. Der ganz überwiegende Teil hatte mehr als ein Jahr Berufserfahrung und arbeitete im stationären Bereich.
b. Thesen zur Zukunft des Gesundheitswesen: Es wurden 11 Thesen aus den teilstrukturierten Interviews formuliert, siehe Tabelle 3 auf Seite 28 dieser Promotionsschrift. Zudem wurden die Thesen auf ihr wahrscheinliches Eintreten gewichtet. Die große Mehrheit der Befragten schätzt auch in Zukunft den Arzt als nicht gegen IT-Systeme austauschbar ein, wenngleich der Einzug neuer Technologien insgesamt zu einem verminderten physischen Kontakt zwischen Arzt und Patient führen wird, und bewertet den persönlichen Arzt-Patienten-Kontakt nach wie vor als relevant, siehe Abbildung 2 auf Seite 29 dieser Promotionsschrift.
c. Auswertung der Lernziele: 288 Lernziele wurden in zwei Befragungsrunden bewertet, davon wurden in der ersten Runde 240 als relevant/sehr relevant (Mean <2,5) und 47 als eher nicht relevant/irrelevant bewertet (Mean >2,5). Ein Lernziel blieb ohne Tendenz. 103 Lernziele kamen anhand der Streuung ihrer Bewertung in die zweite Runde. In diesem Durchgang wurden davon 62 als relevant/sehr relevant und 41 als eher nicht relevant/irrelevant bewertet, d.h. 12 Lernziele wurden abgewertet und vier Lernziele wurden aufgewertet. Insgesamt wurden 231 Lernziele als relevant und 57 als nicht relevant bewertet. Die Relevanz vieler Lernziele der Organsysteme Haut, GI-Trakt, Harn-/Geschlechtsorgane und Sinnesorgane wurde so in Frage gestellt. Auffällig ist hingegen, dass alle Lernziele aus den drei Bereichen Soft Skills, Kommunikation und Psyche als relevant bewertet wurden.

Eine detaillierte Darstellung der oben zusammengefassten Ergebnisse findet sich in [15].

In unserer zweiten Studie fühlte sich der Großteil der 30 teilnehmenden Studierenden zu Beginn des Nachtdienstes unsicher und unvorbereitet, was die Versorgung von Patienten anbelangt

(Mean -0,34). Das Gefühl vorbereitet zu sein stieg bis zum Follow up (Mean 0,66). Vor allem in den Bereichen Anästhesie, Urologie und Anamneseerhebung zeigte sich ein signifikanter Anstieg, siehe Tabelle 2 auf Seite 38 dieser Promotionsschrift. Für das Einschätzen des Gefühls der Selbstsicherheit war die Rolle des Teilnehmenden innerhalb der Gruppe während der Simulation unerheblich. Zudem zeigte sich, dass sowohl die Teammitglieder als auch die Beobachter die Selbstsicherheit des Teamleiters gut einschätzen konnten. Beim Teammitglied ist dies abhängig von seinem eigenen Gefühl der Selbstsicherheit, beim Beobachter ist dies davon unabhängig. Insgesamt wurde die Nachtdienst-Simulation sehr positiv bewertet: von der Betreuung durch die Tutoren, die Möglichkeit Gelerntes anzuwenden bis hin zum Schwierigkeitsgrad der einzelnen Fälle. Besonders das Feedback durch die Gruppentutoren wurde als hilfreich wahrgenommen.
Eine detaillierte Darstellung der oben zusammengefassten Ergebnisse findet sich in [16].

In unserer dritten Studie konnten wir zeigen, dass Studierende, die mit einer Simulationspatientin trainierten, mehr Schamgefühle während des Trainings erfuhren, mehr Zeit mit der Patientin verbrachten sowie mehr pathologische Tastbefunde dokumentierten als die Studierenden, die mit einer Puppe übten, siehe Tabelle 3 auf Seite 46 dieser Promotionsschrift. Schamgefühle waren von der Trainingsmodalität abhängig, und die Unterschiede in der gemessenen Performance korrelierten positiv mit der Abnahme des situativen Schamgefühls. Die Studierenden, die mit einer Simulationspatientin lernten, erfuhren signifikant höhere Schamgefühle während der Intervention als Studierende, die mit einer Puppe übten ($p < 0,001$). Die Abnahme bzw. Zunahme der situativen Schamgefühle beider Interventionsgruppen von der Intervention bis zur Prüfung war signifikant.
Eine detaillierte Darstellung der oben zusammengefassten Ergebnisse findet sich in [20].

d. Diskussion

Bei den für die Delphi-Studie interviewten Experten war eine große Bandbreite an Fachgebieten der Medizin vertreten, dabei hauptsächlich aus dem stationären Bereich und aus großen Kliniken. Unterrepräsentiert waren demnach ambulante Ärzte. Es ist zu bedenken, dass ein Großteil der Zukunftsthesen vor allem die ambulant tätigen Allgemeinmediziner betreffen wird und demnach ihre Einschätzungen von besonderem Interesse wären. Der Fachbereich der Anästhesie, Notfall- und Intensivmedizin ist dagegen überrepräsentiert, was sehr wahrscheinlich damit zusammenhängt, dass in diesem bereits seit längerem viele praktische Übungen in das Medizinstudium integriert sind. Es stellt damit ein Bias dar, welches eventuell die ausgeprägten technischen Trends erklären könnte. Der Verzicht auf eine Bezugnahme zum nationalen OF in den 11 Zukunftsthesen hat methodische Gründe, die in den erheblichen strukturellen und inhaltlichen Unterschieden zwischen dem OF und den Thesen bestehen [2, 3]. Die Kondensation eines Ausbildungsanforderungsprofils auf 11 Thesen machte es im Unterschied zu dem ähnlichen Ansatz der amerikanischen Gruppe [21] möglich, methodisch einen wichtigen Schritt weiter zu kommen, nämlich zu einer Validierung der Thesen. Die Unterscheidung der Lernziele in Kern- und Wahllernziele, wie es die Originalpublikation der OF in Form einer Prioritätsskala zeigt, deckt sich mit den Ergebnissen unserer Delphi-Studie [15]. Auch im Hinblick auf weniger zukunftsrelevante Wahllernziele besteht Übereinstimmung. Die Bewertung der Lernziele, die den Alters-assoziierten Erkrankungen zugeordnet werden können, spiegelt den demographischen Wandel wider. Diese werden von den Experten als relevant eingeordnet und sollten als Kernlernziel behandelt werden, anders als bisher vom NKLM vorgesehen. Trotz der geringeren Repräsentanz der ambulant tätigen Allgemeinmediziner wird die Wichtigkeit der Primärversorgung hervorgehoben, insbesondere im Hinblick auf die Anamnese bei älteren Patienten sowie die Erhebung geriatrischer Scores. Die generierten Thesen bilden die zukünftige Problematik der medizinischen Versorgung außerhalb der großen Zentren in dem Sinne ab, dass sie sowohl eine wachsende Bedeutung von Telemedizin und IT-Systemen als auch die Bedeutsamkeit des Arztes in seiner Person als relevant betonen. Als systemisch allgemein bedeutsam und übergreifend bedeutsam gilt die Entwicklung interprofessioneller Teams und verstärkte Nutzung von IT-Elementen und Strukturen. Die Delphi-Studie zeigt die Zukunftsrelevanz sowohl der großen Kategorien – Kommunikative Fertigkeiten, Soft Skills, Organübergreifende Fertigkeiten, Herz-Kreislauf und Notfall – als auch der kleinen – Psyche und Endokrines System. Bei verschiedenen anderen Organsystemen fällt die Bewertung der fehlenden Zukunftsrelevanz am größten aus: Sinnesorgane, Harn- und Geschlechtsorgane und GI-Trakt. Es fragt sich, ob dieses Ergebnis zum einen durch die Expertenauswahl bedingt ist,

zum anderen auch der Kleinteiligkeit der Lernzielformulierung respektive der verhältnismäßig großen Zahl geschuldet ist. Bei einem Teil der als nicht relevant eingestuften Lernziele, bleibt fraglich, ob dies als generelle Aussage zu werten ist oder die Zuordnung für eine andere Ausbildungsphase vorgezogen würde (zum Beispiel Facharztweiterbildung). Der mögliche Nutzen der explorativen Delphi-Studie liegt darin zu einer Neubewertung des NKLM und damit verbunden eventuell zu einer Überarbeitung zu kommen. Zukünftige Ärzte würden damit auf vielfältigere Weise und differenzierter als bisher auf ihren Berufsstart vorbereitet. Damit würde das weit verbreitete Self-Assessment um die Perspektive der Zukunftsorientierung der gesamten medizinischen Aus- und Weiterbildung ergänzt.

Die Nachtdienst-Studie zeigt ebenso wie bereits vorhandene Studien ein nur gering ausgeprägtes Gefühl des Vorbereitetseins der Medizinstudierenden des letzten Studienjahres [22, 23]. Bereits eine kurze Simulationsphase führte zu einem signifikanten Effekt der Zunahme der Selbstsicherheit [16]. Dabei ist es keine Bedingung, dass jeder Teilnehmende die Simulation aktiv bestreitet sofern er/sie am Prozess der Durchführung aktiv beteiligt ist – mittels Beobachtungs- und Feedbackaufgaben [24, 25]. Bedingung für eine erweiterte Wirksamkeit ist außerdem die systematische Einbeziehung der Beobachter zur Mitbeurteilung der Selbstsicherheit. Rollenwechsel und gelingende Teamarbeit sind wesentliche Bestandteile dieses Simulationslernens [26, 27], was größere Gruppen ermöglicht und die Ausbildungskosten für diese praxisnahe Art des Lernens optimiert. Auffällig in unserer Studie war, dass besonders die Selbstsicherheit im Bereich der Anamneseerhebung zugenommen hat. Worauf dieser beobachtete Effekt ursächlich zurückzuführen ist, kann nicht abschließend erklärt werden. Ein Ansatz für eine Erklärung wäre, dass die Studierenden in jeder der sieben Stationen eine Anamnese erheben mussten, zu der sie jedes Mal direktes Feedback erhielten, was zu einer wachsenden Verfahrenssicherheit führte. Dieser Aspekt der positiven Verstärkung führt von unserer Studie zu den Erkenntnissen von Bloch und Schubert [28, 29], die einen engen Zusammenhang zwischen Selbstsicherheit und einer guten Leistung im Handlungsvollzug (Performance) nachweisen konnten.

Das Konzept der Shame-Studie hat sich in der Praxis eindeutig bewährt: In der Versuchsanordnung wurde großer Wert auf Vergleichbarkeit zwischen den beiden Gruppen gelegt, indem die Unterschiede auf ein Minimum reduziert wurden. Der entscheidende Unterschied lag darin, in der Intervention einmal eine Puppe und zum anderen eine Simulationspatientin einzusetzen. Aus Sicht der CLT ergeben sich folgende Diskussionspunkte:

Die Unterschiede zwischen den Interventionsgruppen bei den Tastbefunden und bei der für die Aufgabe benötigten Zeit können nicht mit Hilfe des situativen Schamgefühls allein erklärt werden. Unsere theoretische Annahme, dass Studierende, die ein stärkeres situatives Schamgefühl erfahren, der Situation eher entkommen möchten und sich daher weniger Zeit nehmen, bestätigte sich nicht unmittelbar in der isolierten Betrachtung der Teilergebnisse. Dafür ist es notwendig, die Ergebnisse in der Zusammenschau zu betrachten und hier ergibt sich ein anderes Bild: Wenn man die Ergebnisse und Qualität der Brustuntersuchung in Relation zur Zu- oder Abnahme des situativen Schamgefühls der jeweiligen Gruppe setzt, zeigt sich eine Gegenläufigkeit zwischen den beiden Gruppen: der signifikanten Zunahme von der Gruppe A steht eine nicht signifikante Abnahme von der Gruppe B gegenüber. Einige Studien zeigten bereits, dass der Einsatz von SPs beim Erlernen von intimen körperlichen Untersuchungen vorteilhaft ist und hilfreich für die Überwindung verhaltenseinschränkender Prädispositionen [30]. Sie sind den praktischen Übung ohne SP überlegen [31]. Mithin ist von einem lohnenden Einsatz von Simulationspatientinnen auszugehen, da Studierende durch ihre Trainingserfahrungen unter Einbezug von Gefühlen [32] einen nachhalten Lerneffekt erfahren. Mit dem Verhalten in der Gruppe B der Studierenden, die Untersuchung zeitintensiver zu gestalten, ist es naheliegend, einen Zusammenhang zu ihrer stärker emotional basierten, intrinsischen Motivation zu sehen. Ob die ausgewählte Versuchsgruppe als repräsentativ gelten kann für zukünftige Studierende der Medizin bleibt eine offene Frage, auch wenn einige Indizes dafür sprechen. Bei der Bewertung der Studierenden in ihrer Prüfung sind zwei Dimensionen zu unterscheiden: die Prozess- und die Ergebnisqualität. Die Versuchsanordnung unserer Studie zeigt keine Unterschiede in der Prozessqualität – bei nur einer OSC-Station zu erwarten – jedoch große Unterschiede bezüglich der Ergebnisse. Es bleibt also die Aufgabe, diesen Aspekt ebenfalls in zukünftige Lehr- und Prüfungssituationen einzubeziehen. Gefühle, die während des Vermittelns praktischer Fertigkeiten hervorgerufen werden, können von großer Bedeutung für die zu lernenden Inhalte sein und die Art des Lernens beeinflussen. Die Fragestellung, wie diese Thematik in der medizinischen Lehre am besten angegangen werden kann, bleibt eine Herausforderung. Simulation erweist sich generell als sinnvolles Format für einige der genannten Herausforderungen. Es ermöglicht vor allem auch das Hervorrufen negativer Emotionen wie beispielsweise das situative Schamgefühl, welches wiederum – richtig eingesetzt – einen guten Lerneffekt zeitigt.

Über den Wert und die Bedeutung der einzelnen Studien hinaus lässt sich allgemein festhalten, dass wichtige fachspezifische Verhaltensdispositionen in wesentlichen Handlungsfeldern

ärztlicher Praxis – sei es an großen Kliniken, sei es in der herkömmlichen Hausarztpraxis – bereits in der Ausbildung auf verschiedenste Weise bearbeitet werden müssen. Die weiterhin vorherrschende, auf kognitives Lernprozesse abzielende Wissensvermittlung bedarf der Zusammenführung von sozioemotionalen und kommunikativen Kompetenzen und der Antizipation zukunftsrelevanter Anforderungen, denen sich die zukünftigen Ärzte werden stellen müssen. Die Weiterentwicklung der medizinischen Lehre im Bereich der praktischen Fertigkeiten unter Berücksichtigung von Emotionen sowie einer gesteigerten Selbstsicherheit kann zudem auch direkte Erfolge bei der Umsetzung eben dieser praktischen Fertigkeiten junger Ärzte zeigen. Für einen effektiven und optimalen Teamprozess ist seine Strukturierung von besonderer Bedeutung. Dabei geht es um die Flexibilisierung und Varianz von Rollenaufgaben für die einzelnen Teilnehmenden ebenso wie um die systematische Nutzung aktiver Beobachtung und von Feedbackverfahren, was die unmittelbare Integration der Beobachterergebnisse zur Selbstsicherheit ermöglicht.

Literaturverzeichnis

1 Ochsmann EB, Zier U, Drexler H, Schmid K. Well prepared for work? Junior doctors' self-assessment after medical education. BMC Med Educ. 2011;11:99. DOI: 10.1186/1472-6920-11-99

2 Hautz SC, Hautz WE, Feufel MA, Spies CD. Comparability of outcome frameworks in medical education: Implications for framework development. Med Teach. 2015;20:1-9. DOI: 10.3109/0142159x.2015.1012490

3 Hautz SC, Hautz WE, Keller N, Feufel MA, Spies C. The scholar role in the National Competence Based Catalogues of Learning Objectives for Undergraduate Medical Education (NKLM) compared to other international frameworks. GMS Ger Med Sci. 2015;13:Doc20. DOI: 10.3205/000224

4 Tokuda Y, Goto E, Otaki J, Jacobs J, Omata F, Obara H, Shapiro M, Soejima K, Ishida Y, Ohde S, Takahashi O, Fukui T. Undergraduate educational environment, perceived preparedness for postgraduate clinical training, and pass rate on the National Medical Licensure Examination in Japan. BMC Med Educ. 2010;10:35.

5 Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol. 2001;52:1–26.

6 Conner M, Norman P: Predicting Health Behaviour. Volume 24. Berkshire, England: Open University Press; 2005.

7 Cortegiani A, Russotto V, Montalto F, Iozzo P, Palmeri C, Raineri SM, Giarratano A. Effect of High-Fidelity Simulation on Medical Students' Knowledge about Advanced Life Support: A Randomized Study. PLoS One. 2015;10:e0125685.

8 Stegmann K, Pilz F, Siebeck M, Fischer F. Vicarious learning during simulations: is it more effective than hands-on training? Med Educ. 2012;46:1001–8.

9 Lindström UH, Hamberg K, Johansson EE. Medical students' experiences of shame in professional enculturation. Medical Education 2011;45:1016–24.

10 McConnell MM, Eva KW. The role of emotion in the learning and transfer of clinical skills and knowledge. Academic Medicine 2012;87:1316–22.

11 Artino AR. When I say ... emotion in medical education. Medical Education 2013;47:1062–3.

12 Bynum WE, Goodie JL. Shame, guilt, and the medical learner: ignored connections and why we should care. Medical Education 2014;48:1045–54.

13 Gilbert, P. Shame, interpersonal behavior, psychopathology, and culture. New York, NY: Oxford University Press, 1998.

14 van Merriënboer JJ, Sweller J. Cognitive load theory in health professional education: Design principles and strategies. Medical Education 2010;44:85–93.
15 Dannenberg KA, Stroben F, Schröder T, Thomas A, Hautz WE. The future of practical skills in undergraduate medical education – an explorative Delphi-Study. GMS Journal for Medical Education 2016;33(4):Doc62.
16 Stroben F, Schröder T, Dannenberg KA, Thomas A, Exadaktylos A, Hautz W. A simulated night shift in the emergency room increases students' self-efficacy independent of role taking over during simulation. BMC Medical Education 2016;16:177.
17 Allen R, Heard J, Savidge M. Global ratings versus checklist scoring in an OSCE. Academic Medicine 1998;73:597–8.
18 Sim JH, Abdul Aziz YF, Vijayanantha A, Mansor A, Vadivelu J, Hassan H. 2015. A closer look at checklist scoring and global rating for four OSCE stations: Do the scores correlate well? Education in Medicine Journal 2015;7:2.
19 Canadian Global Rating Scale. Available at: http://mcc.ca. Accessed April 2, 2014.
20 Hautz WE, Schröder T, Dannenberg KA, März M, Hölzer H, Ahlers O, Thomas A. Shame in Medical Education: A Randomized Study of the Acquisition of Intimate Examination Skills and Its Effect on Subsequent Performance. Teaching and Learning in Medicine 2017;29:196-206.
21 Pershing S, Fuchs VR. Restructuring Medical Education to Meet Current and Future Health Care Needs. Acad Med. 2013;88(12):1798-1801.
22 Cave J, Woolf K, Jones A, Dacre J. Easing the transition from student to doctor: how can medical schools help prepare their graduates for starting work? Med Teach. 2009;31:403–8.
23 Brennan N, Corrigan O, Allard J, Archer J, Barnes R, Bleakley A, Collett T, de Bere SR. The transition from medical student to junior doctor: today's experiences of Tomorrow's Doctors. Med Educ. 2010;44:449–58.
24 Chi MTH, Roy M, Hausmann RGM. Observing tutorial dialogues collaboratively: insights about human tutoring effectiveness from vicarious learning. Cogn Sci. 2008;32:301–41.
25 Grierson LEM, Barry M, Kapralos B, Carnahan H, Dubrowski A. The role of collaborative interactivity in the observational practice of clinical skills. Med Educ. 2012;46:409–16.
26 Hautz WE, Kämmer JE, Schauber S, Spieß CD, Gassmeier W. Diagnostic Performance by Medical Students Working Individually or in Teams. JAMA. 2015;313:303–4.

27 Patterson PD, Pfeiffer AJ, Lave JR, Weaver MD, Abebe K, Krackhardt D, Arnold RM, Yealy DM. How familiar are clinician teammates in the emergency department? Emerg Med J. 2015;32:258–62.

28 Bloch SA, Bloch AJ. Simulation training based on observation with minimal participation improves paediatric emergency medicine knowledge, skills and confidence. Emerg Med J. 2015;32:195–202.

29 Schubert A, Tetzlaff JE, Tan M, Ryckmann V, Mascha E. Consistency, Inter-rater Reliability, and Validity of 441 Consecutive Mock Oral Examinations in Anesthesiology. Anesthesiology. 1999;91:288–98.

30 LeBlanc VR. The effects of acute stress on performance: Implications for health professions education. Academic Medicine 2009;84:S25–33.

31 Hendrickx K, De Winter B, Tjalma W, Avonts D, Peeraer G, Wyndaele JJ. Learning intimate examinations with simulated patients: The evaluation of medical students' performance. Medical Teacher 2009;31:e139–47.

32 Rees CE, Monrouxe LV, McDonald LA. Narrative, emotion and action: Analysing "most memorable" professionalism dilemmas. Medical Education 2013;47:80–96.

2. Eidesstattliche Versicherung

„Ich, Therese Schröder, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: „Verbesserung der medizinischen Lehre durch Einbezug zukunftsrelevanter Praxisinhalte und nicht-kognitiver Faktoren in den medizinischen Lehr-Lernprozess“ selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.
Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren beruhen, sind als solche in korrekter Zitierung (siehe „Uniform Requirements for Manuscripts (URM)“ des ICMJE -*www.**icmje**.org*) kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) entsprechen den URM (s.o) und werden von mir verantwortet.
Meine Anteile an den ausgewählten Publikationen entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Betreuer/in, angegeben sind. Sämtliche Publikationen, die aus dieser Dissertation hervorgegangen sind und bei denen ich Autor bin, entsprechen den URM (s.o) und werden von mir verantwortet.
Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§156,161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

Datum _____ Unterschrift

3. Anteilserklärung an den erfolgten Publikationen

Therese Schröder hatte folgenden Anteil an den folgenden Publikationen:

Publikation 1: Katja A. Dannenberg, Fabian Stroben, **Therese Schröder**, Anke Thomas, Wolf E. Hautz, The future of practical skills in undergraduate medical education – an explorative Delphi-Study, GMS Journal for Medical Education, 2016
Beitrag im Einzelnen: Mitarbeit bei der Datenauswertung, Mitarbeit bei Entwurf und Revision des Manuskriptes, Anfertigung der englischen Übersetzung der Publikation

Publikation 2: Fabian Stroben, **Therese Schröder**, Katja A. Dannenberg, Anke Thomas, Aristomenis Exadaktylos, Wolf E. Hautz, A simulated night shift in the emergency room increases students' self-efficacy independent of role taking over during simulation, BMC Medical Education, 2016
Beitrag im Einzelnen: Mitarbeit bei der Datenerhebung, Mitarbeit bei der Datenauswertung, Mitarbeit bei Entwurf und Revision des Manuskriptes

Publikation 3: Wolf E. Hautz*, **Therese Schröder***, Katja A. Dannenberg, Maren März, Henrike Hölzer, Olaf Ahlers, Anke Thomas, Shame in Medical Education: A Randomized Study of the Acquisition of Intimate Examination Skills and Its Effect on Subsequent Performance, Teaching and Learning in Medicine, 2017
*geteilte Erstautorenschaft
Beitrag im Einzelnen: Konzeption der Studie, Ausarbeitung und Bereitstellung der Lehrmaterialien für die Tutorien, Mitarbeit bei der Datenauswertung, Anfertigung des Manuskriptes, Revision des Manuskriptes

Unterschrift, Datum und Stempel des betreuenden Hochschullehrers/der betreuenden Hochschullehrerin

_____

Unterschrift des Doktoranden/der Doktorandin

_____

Clinical skills OPEN ACCESS This is the original (English) version. The translated (German) version starts at p. 13. article

The future of practical skills in undergraduate medical education – an explorative Delphi-Study

Abstract

Background: 64% of young medical professionals in Germany do not feel adequately prepared for the practical requirements of the medical profession. The goal of "outcome-orientated training" is to structure medical curricula based on the skills needed when entering the workforce after completing undergraduate medical education, and thus to bridge the gap between the skills graduates have attained and those necessary for a career in the medical profession. Outcome frameworks (OFs) are used for this purpose. In preparation for developing the National Competence-Based Catalogue of Learning Objectives for Medicine (NKLM) – the German OF – the "Consensus Statement of Practical Skills in Undergraduate Medical Education" (which structures the teaching and acquisition of practical skills in Germany and which strongly influenced the "Clinical-Practical Skills" chapter of the NKLM) was published in 2011.

It is not uncommon for at least a decade to elapse between the definition and implementation of an OF and the students' graduation, which can further increase the gap between necessary and acquired skills. Thus, the purpose of this paper is to posit theses for future development in healthcare and to apply these theses to a current OF.

Methodology: Partially structured interviews with experts were used to generate theses pertaining to general, future development in healthcare. These theses were assessed by physician experts based on the likelihood of implementation by the year 2025. The 288 learning goals of the consensus statement were assessed for their relevance for medical education in the interim.

Results: 11 theses were generated for the development of medicine, and these theses were assessed and discussed by 738 experts. These theses include the increase in diseases associated with old age, the increasing significance of interprofessional cooperation, and the growing prevalence of telemedicine applications. Of the 288 learning goals of the consensus statement, 231 of the goals were assessed as relevant, and 57 were deemed irrelevant for the short-term future.

Discussion: The theses on the future of healthcare, which were generated in this study and which were validated by numerous experts, provide indications of future developments of overall requirements for medical school graduates. For example, when applied to the content of the "Clinical-Practical Skills" NKLM chapter, they largely validate the future relevance of developing practical skills while also providing indications for their further development as applied to the consensus statement.

Keywords: Skills, practical skills, clinical skills, medical training, consensus method, Delphi survey, learning goals, outcomes, competencies, NKLM

Katja Anne Dannenberg[1,2]
Fabian Stroben[1]
Therese Schröder[3]
Anke Thomas[3]
Wolf E. Hautz[4]

1 Charité – Universitätsmedizin Berlin, Lernzentrum (Skills Lab), Berlin, Germany

2 Charité – Universitätsmedizin Berlin, Department of Emergency Medicine at Campus Benjamin Franklin, Berlin, Germany

3 Charité – Universitätsmedizin Berlin, Gynecology and Obstetrics Clinic, Berlin, Germany

4 Inselspital Bern, University Emergency Center, Bern, Switzerland

Introduction

On the one hand, the significance of obtaining practical skills during undergraduate medical studies has increased significantly in recent years [1], [2]. On the other hand, 64.7% of young medical professionals in Germany state that they do not feel adequately prepared for the practical requirements of the medical profession [3], a figure which is startlingly high, even compared to international data [4], [5]. Possible causes identified by the graduate survey in Cologne (Stosch C et al., unpublished) and a national survey (partially published in [6]), were both the narrow

scope of practical training and the inadequate or lacking integration of this training in curricula and examinations. In order to bridge the gap between education and training, medical curricula are increasingly oriented toward national framework curricula, known as "outcome frameworks" (OF) [7], [8], which – generally speaking – describe the skills and knowledge which should be obtained during a training period in a competence-oriented fashion. Various outcome definitions exist internationally [9], [10], [11]. The Tuning Project [12] in Europe is an attempt to synchronize the many national OFs currently in existence. The German "National Competence-Based Catalogue of Learning Objectives for Medicine" (NKLM) was developed by the medical faculty association in cooperation with the Society for Medical Education (GMA) [13] and was initially published in June of 2015 after a six-year period of development [14]. In preparation for developing the NKLM, the "Consensus Statement of Practical Skills in Undergraduate Medical Education" was developed by the committee for practical skills of the GMA in 2011 [15]. This consensus statement "can and should have a formative effect on faculties to adjust their curricula in accordance with guidelines" [15] and strongly influenced the "Clinical-Practical Skills" chapter of the NKLM. The recommendations of the consensus statement have been implemented and validated within at least one faculty department [16]. In addition, the statement serves to assist the simulator network – a merger of the DACH region Skillslabs – to structure its simulator database [17].
There are, however, notable differences in content and structure between different OFs [18], [19], which raises the question of which OF should reasonably be referenced for teaching proficiency. In addition, developing medical curricula is generally a lengthy process: the six stages of the Kern cycle as a widely taught model of curriculum development [20], for example, require a considerable period of time between the initial definition of requirements, implementation, evaluation, and adaptation. Furthermore, an average of 6.4 years [21] elapse between beginning undergraduate medical education and beginning to practice medicine [21]. This contrasts starkly with rapid developments in medicine and the use of new technologies which have become ubiquitous. Consequently, there is a risk that the contents of curricula developed based on current OFs are no longer up-to-date when the medical professionals educated accordingly enter the medical profession.

1. Object of the Study

The object of this study is to examine the "Consensus Statement of Practical Skills in Undergraduate Medical Education," and thus an important part of the NKLM, for medium-term sustainability. The results should, on the one hand, serve to provide details for the further development of the NKLM; and on the other hand, help enhance the future stability of OF and curricula by means of overarching trends in healthcare which must yet be identified. The applied explorative Delphi method, as well as its results, can also serve to further develop local and national curricula.

2. The explorative Delphi method

Originally developed in the 1950s as a technique for exploring technical developments in a military context [22], this method had been continually developed in the intervening decades [23] and is now considered an established method for analyzing uncertain developments and identifying strategic treatment options [22], [24], [25]. In principle, the Delphi method serves to collect group opinions and to focus group communications [24], as well as to qualitatively and quantitatively assess uncertain facts [24]. Although widely varied definitions of the Delphi method exist [24], certain common basic principles can be identified: anonymity of experts, multiple repetitions of the survey, statistical summary of group opinion, and controlled feedback [22]. The use of the Delphi method has been tested in various contexts, though here it is predominantly of interest to sufficiently documented applications in medical education research, such as for developing guidelines [26], [27], [28], [29].

Methods

The project was structured into preparatory and working phases. During preparations, literary research followed by partially structured stakeholder interviews was used to develop theses about developments in healthcare. These theses were then assessed by means of an expert survey. In addition, the same expert cohort assessed the 288 learning goals of the "Consensus Statement of Practical Skills in Undergraduate Medical Education" within the framework of a 2-level, explorative Delphi survey. The course of the study is depicted in Figure 1.

1. Preparatory Phase: Theses on healthcare development

Guidelines for partially structured interview with various healthcare practitioners were developed by means of selective literature research. The topics discussed in the interview included the following:

- The future development of healthcare
- Potential changes to care and to the disease spectrum
- Changes in medical technology and telemedicine
- Interdisciplinarity and cooperation with other occupational groups
- Future changes to undergraduate medical education
- Medical occupations in Germany and abroad
- Medical skills needed in the future

During the preparatory phase, 9 interviews were conducted with experts in the fields of public health, medical technology and pedagogy, clinical and outpatient, practical skills, and students of human medicine (cf. Table 1for details). Interview partners were chosen by means of a

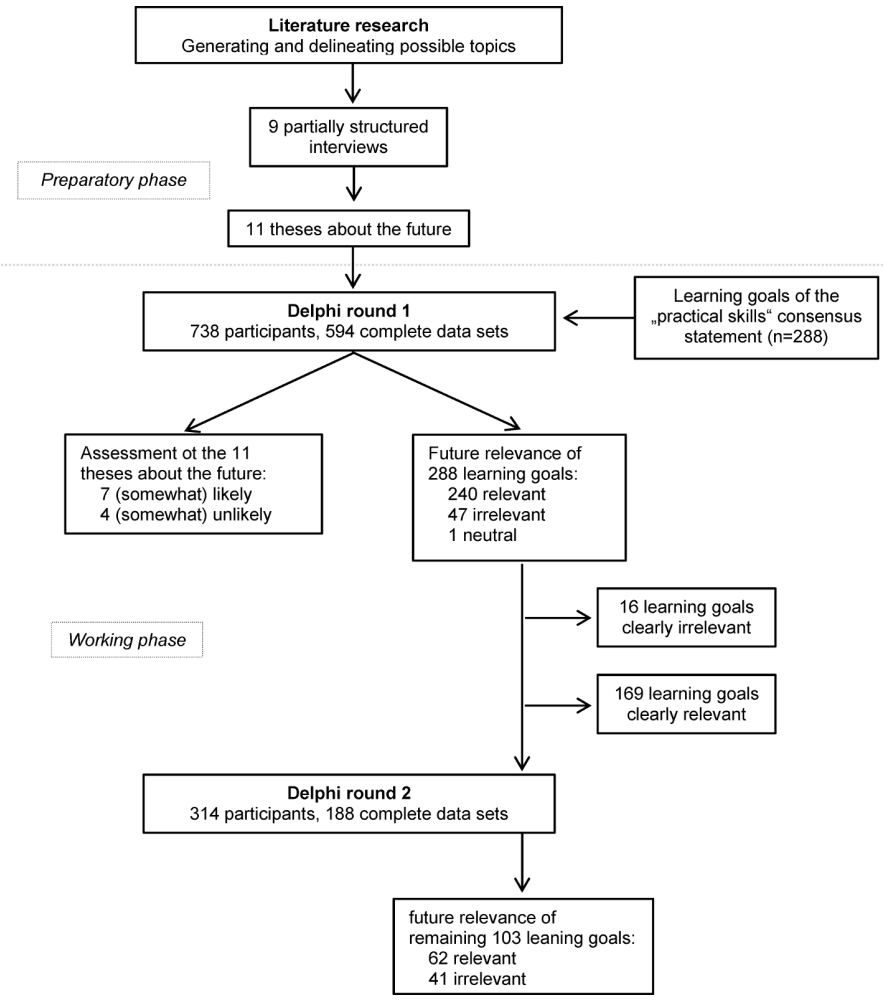

Figure 1: Study design and results overview.

"purposive sampling strategy" [30] with the goal of obtaining as broad a spectrum of perspectives on these topics as possible. All interviewed experts were prepared for the interview. Interviews lasted an average of 22 minutes.

The interviewees' answers were then grouped by topic using a qualitative text analysis and summarized as theses by means of inductive categorization per Mayring [31] by an interdisciplinary research group comprised of two students of futures studies, including one nurse and one student of human medicine with paramedic training, one practicing physician, and one computer scientist. The goal of the Mayring analysis is to systematically process the written communications at hand, and to identify similarities and differences [32]. The principles of categorization were a) category selectivity, and b) a high level of abstraction of the same.

2. Working Phase: Expert interviews on future theses generated, and on consensus statement learning goals

After generating theses, their probability of occurrence was assessed within the framework of an expert interview. The individual learning goals of the "Practical Skills in Undergraduate Medical Education" consensus statement were then assessed by the same experts within the framework of a two-level Delphi study. Physicians in all German medical university hospitals whose email addresses were available on the internet, as well as established physicians, were contacted via email to request their participation in the study. In addition to these 8,000 physicians contacted, others were approached at conferences (e.g., the Skills-Lab Symposium 2012) to request their participation in the study.

Each participant then assessed the probability of occurrence of these theses on the future of medical care (generated during the preparatory phase) using a 4-level Likert scale (1 – very likely to 4 – very unlikely). Afterward, each participant was then assigned randomly to a group of ten in order to assess the future relevance of the consensus statement learning goals. This statement defines 288 learning goals assigned to one of 16 organ systems. There is a statement for three different training stages ("clinical traineeship, practical year, advanced training") based on a three-tiered scale for each learning goal, to what extent this should be mastered ("seen demonstrated, performed under supervision, performed repeatedly"), and the survey further distinguishes between core and elective goals [15]. Each group was asked to assess a portion of the consensus statement learning

goals (ca. 30/group) vis-à-vis their relevance for general medical training up to completing undergraduate medical education in the year 2025 using a 4-level Likert scale (1 – highly relevant to 4 – not at all relevant). The learning goal assessments were depicted based on the degree of mastery required by the advanced training stage, as stipulated in the consensus statement.

After round 1 assessments and individual review of the results by the research group, the round 2 learning goals to be assessed were determined by consensus. Selection criteria included a wide distribution of assessment and the estimated significance of each learning goal as assessed in round 1. The learning goals were then re-evaluated by physicians participating in round 2, who were provided with the result of round 1 evaluations. Registered participants were assigned randomly to two groups for this purpose. Each group re-evaluated circa 50 learning goals.

3. Data evaluation

The online interview was conducted using LimeSurvey (http://www.limesurvey.org). Data evaluation was conducted with Microsoft Excel (Microsoft Corporation, Redmond, WA, USA) and IBM SPSS Statistics 21.0 (SPSS Inc., Chicago, IL, USA).

Results

1. Participants

738 experts registered online for round 1 of the study, and 594 complete data sets (19.5% dropout rate) were usable. 314 experts were registered for round 2, and 188 complete data sets (40.1% dropout rate) were usable. Since the learning goals were assessed within the context of the organ system assigned to them, we only took complete data sets into consideration.

Participants were only asked to assess theses about the future during round 1. Partially-completed questionnaires were taken into consideration here, as the theses about the future can logically be interpreted on an individual basis. For this purpose, 651 expert opinions (11.8% dropout rate) were available.

Table 1: Overview of interview partners

Interview participant	1	2	3	4	5	6	7	8	9
Work environment	Rural area	University	University	University	University	Professional association	University	Urban area	Rural area / Scandinavia
Profession	General medicine	Public Health/ Health Science	Physician, Master of Medical Education	Student of human medicine	Resident physician/ employed in a teaching organization	Medical technology	Nursing education	Public Health/ Health science	Emigrated specialist
Sex	M	M	M	F	F	M	M	M	F
Interview duration in min	18:44	28:11	24:34	24:14	17:20	18:04	28:31	19:59	16:45

A large portion of experts in round 1 had more than one year of work experience (96.0%), and 137 experts had been practicing for more than 15 years. The overwhelming majority of physicians worked on an inpatient basis in maximum-care hospitals (87.9%). Study participants represented a total of 26 disciplines. The number of experts in round 2 was smaller, though the characteristics of their working environments were similar. In comparison to German Medical Association (BÄK) statistics from 2014, the proportion of inpatient physicians is large (88.6% in round 1 and 79.3% in round 2, compared to 51.0% by BÄK figures), and the exact distribution of specialties also differed somewhat. Overall, 57.8% (round 1) and 53.7% (round 2) of experts work in disciplines such as surgery, internal medicine, or anesthesia, in addition to general medicine. This is consistent with BÄK data, which lists this figure at 48.8% [http://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/Statistik2014/Stat14AbbTab.pdf]. An exact analysis of the participant cohort in comparison to BÄK figures from 2014 is depicted in Table 2.

2. Theses on the future of healthcare

A total of 11 theses on the future of health care were derived from the partially structured interviews (cf. Table 3).

The theses on the future of healthcare trends generated in the preparatory phase are listed below. It was assumed that these theses were listed in accordance with expert assessment (cf. Table 3 and Figure 2):

Aspects of managing dementia are becoming significantly more important in physician communication. Increasingly balanced patient-physician relationships are emphasizing non-authoritarian forms of discussion and reasoning.

Increasing mechanization also poses a barrier to entry in the medical field: the relevance of purely manual skills is decreasing, yet IT technologies still cannot determine medical history or diagnoses, requiring the skills of doctors. Diagnostics and patient monitoring will lead to less physical contact, and patients prefer internet and smartphones for this purpose.

The physician remains the personal point of contact in established practice concepts. Duties once performed purely by physicians are, however, increasingly being delegated or substituted. Mobile treatment concepts of primary care are not catching on.

Business economic considerations are slowly moving into focus: while business economic and organizational aspects are included in training, patients' financial concerns also play a role in the type of care and treatment.

3. Assessment of learning goals

288 learning goals were assessed by experts in the 2 rounds of the study. The average of all expert assessments was used to determine whether a learning goals was deemed relevant or irrelevant. In round 1 of the Delphi study, 240 learning goals were assessed as relevant or highly relevant (average<2.5) and 57 were assessed as somewhat relevant or irrelevant (average>2.5), while 1 learning goal was assessed as neither (average=2.5).

After reviewing the results of round 1, 103 learning goals were selected for assessment in round 2 based on the distribution of round 1 results; of these, 71 were assessed as relevant, 31 were assessed as irrelevant, and 1 was assessed as neutral in round 1. In round 2, experts assessed 62 learning goals as relevant and 41 as irrelevant. In comparing the two rounds, 13 learning goals (12.6%) were rated less relevant and 4 (3.9%) were rated as more relevant in round 2. A total of 231 learning goals were considered relevant and 57 learning goals were considered irrelevant. Figure 1 depicts an overview of these results.

A more extensive review based on organ systems revealed that a large portion of learning goals for the sensory organ system (65.0%) were assessed as irrelevant for the future. Likewise, the future relevance of numerous learning goals associated with the skin, urogenital, and GI tract organs (≥30.0% each) was called into question.

31 of the 55 (56.4%) elective learning goals were assessed as irrelevant for the future – on the other hand, more than half (54.4%) of non-relevant learning goals are elective. Only 8% of the learning goals and skills which should be mastered upon completing advanced training were assessed as irrelevant for the future, yet 42% of skills were assessed at the lowest level of skill ("seen demonstrated"). The exact results are depicted in Table 4.

The online appendix of this study shows an overview of all consensus statement learning goals and their assessments in the two rounds of the Delphi study.

Discussion

This paper attempts, on the one hand, to anticipate future global requirements for medical school graduates; and on the other, to concretely examine the future relevance of practical medical skills as an example of a limited scope of competency in undergraduate medical studies. For the latter, the learning goals of the consensus statement of practical skills in undergraduate medical studies was assessed by means of an explorative Delphi study as preliminary work to the NKLM "clinical-practical skills" chapter.

The panel of experts in the Delphi study possesses many years of work experience and represents nearly all medical specialties. The majority of experts work in maximum-care hospitals, including university hospitals in which undergraduate training is primarily conducted in Germany. Though the experts' more intensive knowledge has influenced the contents and requirements of the study, outpatient physicians (who are needed primarily for contributing opinions on these theses about the future of medical care) are underrepresented. We can only speculate

Table 2: Work experience, environment, and specialties for participants in rounds 1 and 2.

	Delphi round 1		Delphi round 2		Practicing physicians in GER in 2014	
	n	%	n	%	n	%
total	594	100.0	188	100.0	365 247	100.0
Time elapsed since receiving license to practice medicine:						
< 1 year	19	3.2	4	2.1		
1–5 years	186	31.3	55	29.3		
6–10 years	147	24.7	41	21.8		
11–15 years	100	16.8	34	18.1		
> 15 years	137	23.1	46	24.5		
No response	5	0.8	8	4.3 *		
Primary working environment						
Inpatient/maximum care	522	87.9	146	77.7	186 329	51.0
Inpatient/standard and primary care	4	0.7	3	1.6 *		
Outpatient/urban	15	2.5	7	3.7	147 948	40.5 *
Outpatient/rural	10	1.7	7	3.7 *		
Medical education research	26	4.4	11	5.9	30 970	8.5
Other & no response	17	2.9	14	7.4 *		
Primary specialty					*N=255 020*[a]	
Anesthesia and intensive medicine	139	23.4	45	23.9	22 071	8.7 *
Internal medicine	118	19.9	32	17.0	60 697	23.8
Surgery	86	14.5	24	12.8	41 544	16.3
Pediatrics, gynecology, obstetrics	78	13.1	28	14.9	33 408	13.1
ENT & ophthalmology	55	9.3	10	5.3	13 239	5.2
Neurological disciplines	44	7.4	8	4.3	21 208	8.3
Radiological disciplines	24	4.0	6	3.2	10 099	4.0
General medicine	17	2.9	14	7.4 *	43 206	16.9 *
Other & no response	33	5.5	21	11.2 *	9 548	3.7

Notes: Percents based on total amount. Percents marked with * in round 2 or BÄK 2014 deviate by a minimum of 50% from those in round 1. [a] Physicians in advanced training are not listed in BÄK reviews with their desired training, and are thus not included here.
Surgery also includes urology, maxillofacial surgery; radiological disciplines also include radiology, radiation therapy, and nuclear medicine; neurological disciplines also include neurology, neurological medicine, neurosurgery, and psychiatry; pediatrics, gynecology, and obstetrics also include child and adolescent psychiatry; internal medicine also includes psychosomatic medicine, dermatology, physical medicine, and rehabilitation.

on the reasons for which so few outpatient physicians participated in the study.

At the beginning of this paper, 11 theses for relevant topics for future medical education were identified. We deliberately chose not to provide an OF as a basis for the interviews, as national OFs differ substantially in structure [18] as well as content [19].

An American group has already published a similar approach [33] which did not, however, take any further validation steps for its theses in comparison to our study. Below, we will discuss a few theses of this study in the context of their assessment by participating experts and derive implications for medical education:

The current demographic shift has caused an increase in diseases associated with advanced age, such as mild cognitive impairment and dementia [34], [35]. The learning goals catalog lists two learning goals which could be attributed to dementia illnesses. Determining the medical history of elderly patients and performing simple test procedures such as geriatric assessments or falling risk tests were both assessed as relevant for the future. However, emphasizing geriatric test procedures was listed in the catalog as an elective goal, though it should be a core learning goals according to the experts in this study.

The state of medical care, especially in rural areas with their own specific requirements [36], is in need of improvement [37]. Experience obtained in these places during voluntary training during undergraduate medical studies seemed to have a positive effect on students' learning and career choices [38] and could improve primary care. The mandatory primary care physician clinical traineeship [39] was recently introduced, and a GMA position paper emphasizes the significance of primary care during undergraduate medical studies [40].

In addition to these structural changes, working on an (interprofessional) team and using telemedicine or E-health applications will become more important in the future: delegating work to non-medical personnel increases the effectiveness of primary care [41], [42]. At the same time, learning goals which do not apply solely to physicians, such as applying plaster casts or demonstrating functional taping, have also been assessed as relevant for the future. Beginning to promote interprofessional cooperation during undergraduate medical studies, such as combined courses with trainees or students of other medical care professions, could be one method of implementing the interpersonal aspect of care more intensively during medical education.

Table 3: Expert responses (round 1) to the 11 stakeholder theses

Theses	N	Mean	SD	Very likely	Somewhat likely	Somewhat unlikely	Very unlikely	No response
1. Based on the demographic shift, special knowledge and skills in managing patients with dementia with be of great significance for all specialties in adult medicine in the year 2025.	626	1.65	0.69	278	273	46	11	18
2. Remote patient monitoring, video telephone services, and transmitting laboratory results using the internet and smartphones are accepted in the year 2025, and these services are used for the majority of patients.	626	1.85	0.83	244	232	112	21	17
3. In the year 2025, procedures performed previously only by physicians will also be performed and assessed by non-physician personnel.	651	1.85	0.91	277	232	83	48	11
4. A doctor in the year 2025 is a health manager whose training and education in basic knowledge of organization and business administration must be enhanced.	651	1.90	0.79	209	313	83	30	16
5. New diagnostic and treatment options due to innovative IT applications will lead to less physical contact between treating physicians and patients in the future.	651	2.13	0.84	146	303	137	46	19
6. The information gap between physician and patient will continue to decrease. Consequently, it is not the physician's authority that makes final decisions regarding treatment, but rather his or her ability to communicate and reason with the patient regarding diagnosis and treatment.	626	2.18	0.77	117	285	174	26	24
7. Financial concerns are the primary criteria for selecting supply and treatment options for patients in 2025.	651	2.25	0.88	142	243	199	50	17
8. The physician's most important tools in 2025 are his hands.	651	2.83	0.96	83	113	251	174	30
9. In 2025, primary medical care with primary care physicians and specialists is guaranteed overwhelmingly by means of mobile treatment plans such as daily practices, house visits, or buses, rather than in established local practices.	626	2.90	0.80	33	131	294	139	29
10. Medical history and diagnosis will be established automatically by certified IT systems. If needed, specially trained physicians will be consulted.	626	2.94	0.82	35	122	293	158	18
11. In 2025, the physician, in his person, is interchangeable, and is chosen by patients based on his function: access to treatment and diagnosis, rather than personal contact, is the deciding factor.	651	3.03	0.92	48	123	221	238	21

Notes: Assessed with 4-tier Likert scale (1 = very likely […] 4 = very unlikely)

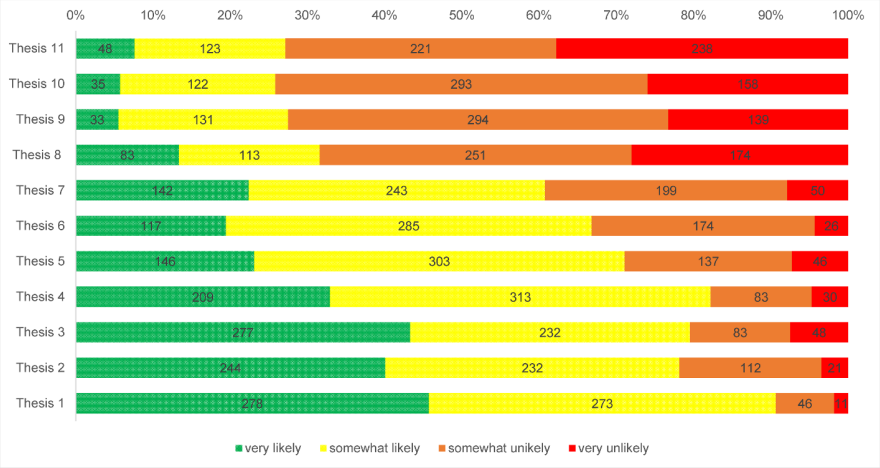

Figure 2: Graphical depiction of the experts' assessment of the 11 theses on the future of healthcare. See tab. 3 for allocation of these theses.

Table 4: Results of Delphi Rounds 1 + 2

	total	Relevant	Irrelevant	
	n	n	n	%
total	288	231	57	19.8
Organ Systems				
Sensory organs	40	14	26	65.0
Urogenital	23	15	8	34.8
GI tract	12	8	4	33.3
Skin	13	9	4	30.8
Respiratory	10	8	2	20.0
Nervous system	10	8	2	20.0
Musculoskeletal	11	9	2	18.2
Growth/aging	21	18	3	14.3
Blood/immune	10	9	1	10.0
Cardiovascular	22	20	2	9.1
Border area, emergency	36	34	2	5.6
Spanning multiple organ systems	36	35	1	2.8
Border area, soft skills	24	24	0	0.0
Border area, communication	14	14	0	0.0
Psyche	4	4	0	0.0
Endocrine system	2	2	0	0.0
Elective/core learning goals				
Elective learning goals	55	24	31	56.4
Core learning goals	232	206	26	11.2
Core/elective learning goals	1	1	0	0.0
Extent of mastery by completion of advanced training, per the consensus statement				
Seen demonstrated	62	36	26	41.9
Performed a few times	61	44	17	27.9
Performed routinely	165	151	14	8.5

Notes: Table organized by organ system, elective/core learning goals, and extent of mastery, as determined in consensus statement. Percent values calculated line by line.

In addition, electronic support system (health information technology) resources that are currently available could be used more effectively [42], [43]. Some positive effects of this technology, such as increased activity for COPD patients [44] or improved control of chronic asthma symptoms [45], have already been demonstrated. Integrating this growing field into training and education seems crucial, and could take place by means of telemedicine modules [46]. A dedicated learning goals catalog for e-health and telemedicine has already been published and can be consulted for future developments [47].

In consideration of this knowledge and the high relevance of soft skills and communication ability in this Delphi study, telephone- or internet-based physician-patient interaction could also grow more significant [48]. In order to do this field justice, training for communication skills (e.g., via telephone) should be intensified [49], as has been implemented in individual cases [50], [51]. Older patients in Germany see telemedicine methods more critically, however, and miss personal contact with and direct feedback from their physician [52].

Prioritization of core and elective learning goals in the OF original publication [15] (which were, in part, determined by means of the Delphi methodology and our results) mutually validate each other. More than 90% of the consensus statement learning goals assessed as needing to be mastered and nearly all border area learning goals were assessed by our participants as especially relevant

for the future of the medical profession. On the other hand, more than 50% of skills listed in the consensus statement as elective were also assessed as less relevant by our study.

The experts assessed practical learning goals overwhelmingly as relevant for the future, primarily in the large categories of communication skills, soft skills, interinstitutional skills, cardiovascular, and emergency, but also in narrower disciplines such as mental health and the endocrine system. The portion of learning goals not relevant for the future is largest for the sensory organs and for the urogenital and GI tract systems. This could be related to the choice of experts, but could also be due to the fact that the learning goals were phrased in a very specific manner, and thus there are a great many of them. In other catalog categories, more learning goals tended to be summarized as one, which made it difficult for experts to provide a differentiated assessment. In addition, it cannot be determined whether rejection of a learning goal was due to a general lack of future relevance, or because experts considered the learning goal relevant for the future, but believed that the learning goal should be a part of specialty training rather than general medical training.

In a further step, the detailed results of this study (cf. online appendix) could be used, just like other validation studies [16], to re-assess the individual learning goals of the consensus statement and the NKLM, which would contribute to a review of the consensus statement and the NKLM.

Preparing future physicians to practice medicine should be done, in parallel, on as many levels as possible. Numerous OFs, unlike this study, currently name the "self-directed learning" method resulting from "self-assessment" [53] as a significant method of improving the results of undergraduate medical studies [54]. At the same time, there are significant doubts as to the accuracy of self-assessments [53], [55], [56]. Thus, "lifelong learning" based on self-assessment cannot be the only methods of anticipating and addressing future developments.

This should be done at the level of the OF. In addition to instructions on effective, self-directed, and lifelong learning, optimizing current OFs could make significant contributions. Anticipating future developments in conjunction with current research results could, on the one hand, provide important incentives for new content, and on the other hand, explorative Delphi studies could examine current learning goals and OFs with respect to their sustainability and possibly identify deficiencies. In concrete terms, the results of the Delphi study could serve to justify specific revisions to and implementation of the NKLM in different departments. Broad trend-setting decisions on possible future trends in undergraduate medical education can be derived from the assessment of these theses on the future.

1. Limitations

Cognitive bias must be accepted as a significant limitation of any expert survey. This is of particular importance for the application of the explorative Delphi method for assessing issues that are uncertain, per se, as they take place in the future, since the line between rational assessment and the experts' personal desires or fears could be blurred [57]. The structure of the survey could also have influenced expert opinions. After first assessing the likelihood of implementing theses in the future, experts were then asked to assess the future relevance of learning goals. This could have led to a bias. Though experts were asked to base their assessments on general education leading up to the medical exam, it cannot be determined whether this was actually done, and to what extent experts assessed general education as opposed to specialty training.

The study population of this paper is comprised primarily of inpatient physicians at maximum-care facilities. Consequently, a bias against outpatient treatment methods cannot be ruled out. In addition, colleagues practicing general medicine were underrepresented (just 2.9% of experts), while anesthesiology and intensive care – two highly specialized subjects – were overrepresented, which could explain the strong emphasis on technological trends in the theses generated. One possible explanation for the large number of anesthesiology practitioners represented in the study could be their disproportionate involvement in conveying practical skills. A follow-up survey with primarily outpatient physician seems reasonable.

Conclusions

The explorative Delphi method provides an adequate opportunity to allow a current outcome framework to be assessed by experts on the basis of its future relevance. In addition, future trends can be anticipated by means of generating and assessing theses. It is important to continually review and adapt current OF and curricula to future developments in order to provide optimal preparation for medical studies graduates for their future daily professional lives.

Acknowledgements

We would like to thank Sascha Dannenberg for his help and support with technical questions about futures studies, and for his assistance in carrying out the study. In addition, we would like to thank Nicole Ambacher and Daniel Knapp for their support in data collection. Finally, we would like to sincerely thank all of the experts who took part in this study.

Data

Data for this article are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.q4sc8 [58]

Competing interests

The authors declare that they have no competing interests.

Attachments

Available from
http://www.egms.de/en/journals/zma/2016-33/zma001061.shtml

1. Anhang.pdf (183 KB)
   Online attachment – only in german

References

1. Nicky Hudson J, Tonkin AL. Clinical skills education: outcomes of relationships between junior medical students, senior peers and simulated patients. Med Educ. 2008;42(9):901-908. DOI: 10.1111/j.1365-2923.2008.03107.x
2. Yudkowsky R, Otaki J, Lowenstein T, Riddle J, Nishigori H, Bordage G. A hypothesis-driven physical examination learning and assessment procedure for medical students: initial validity evidence. Med Educ. 2009;43(8):729-740. DOI: 10.1111/j.1365-2923.2009.03379.x
3. Ochsmann EB, Zier U, Drexler H, Schmid K. Well prepared for work? Junior doctors' self-assessment after medical education. BMC Med Educ. 2011;11:99. DOI: 10.1186/1472-6920-11-99
4. Cave J, Woolf K, Jones A, Dacre J. Easing the transition from student to doctor: how can medical schools help prepare their graduates for starting work? Med Teach. 2009;31(5):403-408. DOI: 10.1080/01421590802348127
5. Goldacre MJ, Taylor K, Lambert TW. Views of junior doctors about whether their medical school prepared them well for work: questionnaire surveys. BMC Med Educ. 2010;10:78. DOI: 10.1186/1472-6920-10-78
6. Damanakis A. Ein kleiner Statusreport zu Skills Labs in der D-A-CH-Region. Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA). München, 05.-08.10.2011. Düsseldorf: German Medical Science GMS Publishing House; 2011. Doc11gma133. DOI: 10.3205/11gma133
7. Harris P, Snell L, Talbot M, Harden RM. Competency-based medical education: implications for undergraduate programs. Med Teach. 2010;32(8):646-650. DOI: 10.3109/0142159X.2010.500703
8. Busing N, Slade S, Rosenfield J, Gold I, Maskill S. In the spirit of Flexner: working toward a collective vision for the future of medical education in Canada. Acad Med. 2010 Feb;85(2):340-348. DOI: 10.1097/ACM.0b013e3181c8880d
9. Bloch R, Burgi H. The Swiss catalogue of learning objectives. Med Teach. 2002 Mar;24(2):144-150. DOI: 10.1080/01421590220120759
10. Frank JR, Danoff D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. Med Teach. 2007;29(7):642-647. DOI: 10.1080/01421590701746983
11. Metz JC. 'Blueprint 1994': common objectives of medical education in The Netherlands. Neth J Med. 1999;55(4):165-167. DOI: 10.1016/S0300-2977(99)00067-4
12. Cumming A, Ross M. The Tuning Project for Medicine-learning outcomes for undergraduate medical education in Europe. Med Teach. 2007;29(7):636-641. DOI: 10.1080/01421590701721721
13. Hahn EG, Fischer MR. Nationaler Kompetenzbasierter Lernzielkatalog Medizin (NKLM) für Deutschland: Zusammenarbeit der Gesellschaft für Medizinische Ausbildung (GMA) und des Medizinischen Fakultätentages (MFT). GMS Z Med Ausbild. 2009;26(3):Doc35. DOI: 10.3205/zma000627
14. Fischer MR, Bauer D, Mohn K, NKLM-Projektgruppe. Finally finished! National Competence Based Catalogues of Learning Objectives for Undergraduate Medical Education (NKLM) and Dental Education (NKLZ) ready for trial. GMS Z Med Ausbild. 2015;32(3):Doc35. DOI: 10.3205/zma000977
15. Schnabel KP, Boldt PD, Breuer G, Fichtner A, Karsten G, Kujumdshiev S, Schmidts M, Stosch C. A consensus statement on practical skills in medical school - a position paper by the GMA Committee on Practical Skills. GMS Z Med Ausbild. 2011;28(4):Doc58. DOI: 10.3205/zma000770
16. Blaum WE, Dannenberg KA, Friedrich T, Jarczewski A, Reinsch AK, Ahlers O. The practial use of the consensus statement on practical skills in medical school-a validation study. GMS Z Med Ausbild. 2012;29(4):Doc58. DOI: 10.3205/zma000828
17. Damanakis A, Blaum WE, Stosch C, Lauener H, Richter S, Schnabel KP. Simulator Network project report: a tool for improvement of teaching materials and targeted resource usage in Skills Labs. GMS Z Med Ausbild. 2013;30(1):Doc4. DOI: 10.3205/zma000847
18. Hautz SC, Hautz WE, Feufel MA, Spies CD. Comparability of outcome frameworks in medical education: Implications for framework development. Med Teach. 2015;20:1-9. DOI: 10.3109/0142159x.2015.1012490
19. Hautz SC, Hautz WE, Keller N, Feufel MA, Spies C. The scholar role in the National Competence Based Catalogues of Learning Objectives for Undergraduate Medical Education (NKLM) compared to other international frameworks. GMS Ger Med Sci. 2015;13:Doc20. DOI: 10.3205/000224
20. Kern DE. Curriculum development for medical education: a six-step approach. Baltimore u.a.: Johns Hopkins Univ. Press; 1998.
21. Wijnen-Meijer M, Burdick W, Alofs L, Burgers C, ten Cate O. Stages and transitions in medical education around the world: clarifying structures and terminology. Med Teach. 2013;35(4):301-307. DOI: 10.3109/0142159X.2012.746449
22. Rowe G, Wright G. The Delphi technique as a forecasting tool: issues and analysis. I J Forecast. 1999;15(4):353-375. DOI: 10.1016/S0169-2070(99)00018-7
23. Grunwald A. Modes of orientation provided by futures studies: making sense of diversity and divergence. Eur J Futures Res. 2013;2(1).
24. Häder M. Delphi-Befragungen: ein Arbeitsbuch. 2. Aufl. ed. Wiesbaden: VS; 2009. DOI: 10.1007/978-3-531-91926-3
25. Gönder D. KiTA 2030: Eine Delphi-Befragung zur Zukunft von Kindertageseinrichtungen in Deutschland. iF-Schriftenreihe Sozialwissenschaftliche Zukunftsforschung 04/2014. Berlin: Freie Universität Berlin, Institut Futur; 2014.
26. Williams PL, Webb C. The Delphi technique: a methodological discussion. J Adv Nurs. 1994;19(1):180-186. DOI: 10.1111/j.1365-2648.1994.tb01066.x

27. Houwink EJ, Henneman L, Westerneng M, van Luijk SJ, Cornel MC, Dinant JG, Vleuten Cv. Prioritization of future genetics education for general practitioners: a Delphi study. Genet Med. 2012;14(3):323-329. DOI: 10.1038/gim.2011.15

28. De Villiers MR, de Villiers PJ, Kent AP. The Delphi technique in health sciences education research. Med Teach. 2005;27(7):639-643. DOI: 10.1080/13611260500069947

29. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs. 2000;32(4):1008-1015.

30. Tavakol M, Sandars J. Quantitative and qualitative methods in medical education research: AMEE Guide No 90: Part II. Med Teach. 2014;36(10):838-848. DOI: 10.3109/0142159X.2014.915297

31. Mayring, P. Qualitative Inhaltsanalyse. Grundlagen und Techniken. 11., aktualisierte und überarbeitete Auflage, Weinheim, Basel: Beltz Verlag; 2010.

32. Flick U, von Kardorff E, Steinke I. Qualitative Forschung. Ein Handbuch. Orig.-Ausg., 8. Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verl; 2010.

33. Pershing S, Fuchs VR. Restructuring Medical Education to Meet Current and Future Health Care Needs. Acad Med. 2013;88(12):1798-1801. DOI: 10.1097/ACM.0000000000000020

34. Pater C. Mild cognitive impairment (MCI) - the novel trend of targeting Alzheimer's disease in its early stages - methodological considerations. Curr Alzheimer Res. 2011;8(7):798-807. DOI: 10.2174/156720511797633250

35. Langa KM, Levine DA. The diagnosis and management of mild cognitive impairment: a clinical review. JAMA. 201417;312(23):2551-2561.

36. Weeks WB, Wallace AE. Rural-urban differences in primary care physicians' practice patterns, characteristics, and incomes. J Rural Health. 2008;24(2):161-170. DOI: 10.1111/j.1748-0361.2008.00153.x

37. Rosenblatt RA, Andrilla CH, Curtin T, Hart LG. Shortages of medical personnel at community health centers: implications for planned expansion. JAMA. 2006;295(9):1042-1049. DOI: 10.1001/jama.295.9.1042

38. Barrett FA, Lipsky MS, Lutfiyya MN. The impact of rural training experiences on medical students: a critical review. Acad Med. 2011;86(2):259-263. DOI: 10.1097/ACM.0b013e3182046387

39. Bundesministerium für Gesundheit. Erste Verordnung zur Änderung der Approbationsordnung für Ärzte vom 17. Juli 2012. Bundesgesetzbl. 2012;I(34);1539. Zugänglich unter/available from: http://www.bgbl.de/xaver/bgbl/start.xav?start=%2F%2F*[%40attr_id%3D%27bgbl112s1539.pdf%27]#__bgbl__%2F%2F*[%40attr_id%3D%27bgbl112s1539.pdf%27]__1436264296102

40. Huenges B, Gulich M, Böhme K, Fehr F, Streitlein-Böhme I, Rüttermann V, Baum E, Niebling WB, Rusche H. Recommendations for undergraduate training in the primary care sector-position paper of the GMA-Primary Care Committee. GMS Z Med Ausbild. 2014;31(3):Doc35. DOI: 10.3205/zma000927

41. Altschuler J, Margolius D, Bodenheimer T, Grumbach K. Estimating a reasonable patient panel size for primary care physicians with team-based task delegation. Ann Fam Med. 2012;10(5):396-400. DOI: 10.1370/afm.1400

42. Chen PG, Mehrotra A, Auerbach DI. Do we really need more physicians? Responses to predicted primary care physician shortages. Med Care. 2014;52(2):95-96. DOI: 10.1097/MLR.0000000000000046

43. Urban E, Ose D, Joos S, Szecsenyi J, Miksch A. Technical support and delegation to practice staff - status quo and (possible) future perspectives for primary health care in Germany. BMC Med Inform Decis Mak. 2012;12:81. DOI: 10.1186/1472-6947-12-81

44. Lundell S, Holmner A, Rehn B, Nyberg A, Wadell K. Telehealthcare in COPD: a systematic review and meta-analysis on physical outcomes and dyspnea. Respir Med. 2015;109(1):11-26. DOI: 10.1016/j.rmed.2014.10.008

45. Fathima M, Peiris D, Naik-Panvelkar P, Saini B, Armour CL. Effectiveness of computerized clinical decision support systems for asthma and chronic obstructive pulmonary disease in primary care: a systematic review. BMC Pulm Med. 2014;14:189. DOI: 10.1186/1471-2466-14-189

46. Brockes MC, Wirth F, Schmidt-Weitmann S, Battegay E, Gerke W. Implementierung des Moduls Klinische Telemedizin/E-Health" in das Medizinstudium der Universität Zürich. GMS Z Med Ausbild. 2010;27(1):Doc14. DOI: 10.3205/zma000651

47. Rohrig R, Stausberg J, Dugas M, GMDS project group Medical Informatics Education in Medicine". Development of national competency-based learning objectives "Medical Informatics" for undergraduate medical education. Methods Inf Med. 2013;52(3):184-188. DOI: 10.3414/ME13-04-0001

48. Car J, Sheikh A. Telephone consultations. BMJ. 2003;326(7396):966-969. DOI: 10.1136/bmj.326.7396.966

49. Car J, Freeman GK, Partridge MR, Sheikh A. Improving quality and safety of telephone based delivery of care: teaching telephone consultation skills. Qual Saf Health Care. 2004;13(1):2-3. DOI: 10.1136/qshc.2003.009241

50. Fisher J, Martin R, Tate D. Hands on + hands free: simulated on-call interaction. Clin Teach. 2014;11(6):425-428. DOI: 10.1111/tct.12180

51. Dickinson M, Pimblett M, Hanson J, Davis M. Reflecting reality: pager simulations in undergraduate education. Clin Teach. 2014;11(6):421-424. DOI: 10.1111/tct.12185

52. Terschuren C, Mensing M, Mekel OC. Is telemonitoring an option against shortage of physicians in rural regions? Attitude towards telemedical devices in the North Rhine-Westphalian health survey, Germany. BMC Health Serv Res. 2012;12:95. DOI: 10.1186/1472-6963-12-95

53. Eva KW, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. Acad Med. 2005;80(10 Suppl):S46-54. DOI: 10.1097/00001888-200510001-00015

54. Hautz SC, Hautz WE, Feufel MA, Spies CD. What makes a doctor a scholar: a systematic review and content analysis of outcome frameworks. BMC Med Educ. 2015. Under Review.

55. Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. JAMA. 2006;296(9):1094-1102. DOI: 10.1001/jama.296.9.1094

56. Li ST, Paterniti DA, Co JP, West DC. Successful self-directed lifelong learning in medicine: a conceptual model derived from qualitative analysis of a national survey of pediatric residents. Acad Med. 2010;85(7):1229-1236. DOI: 10.1097/ACM.0b013e3181e1931c

57. Gerhold L. Standards und Gütekriterien der Zukunftsforschung: ein Handbuch für Wissenschaft und Praxis. Wiesbaden: Springer VS; 2015. DOI: 10.1007/978-3-658-07363-3

58. Dannenberg KA, Stroben F, Schröder T, Thomas A, Hautz WE. Data from: The future of practical skills in undergraduate medical education - an explorative Delphi-Study. Dryad Digital Repository. 2016. DOI: 10.5061/dryad.q4sc8

Corresponding author:
Katja Anne Dannenberg
Charité – Universitätsmedizin Berlin, Lernzentrum (Skills Lab), Charitéplatz 1, D-10117 Berlin, Germany, Phone: +49 (0)30/450-576403, Fax: +49 (0)30/450-576922
katja-anne.dannenberg@charite.de

Please cite as
Dannenberg KA, Stroben F, Schröder T, Thomas A, Hautz WE. The future of practical skills in undergraduate medical education – an explorative Delphi-Study. GMS J Med Educ. 2016;33(4):Doc62.
DOI: 10.3205/zma001061, URN: urn:nbn:de:0183-zma0010614

This article is freely available from
http://www.egms.de/en/journals/zma/2016-33/zma001061.shtml

Received: *2015-06-25*
Revised: *2016-02-29*
Accepted: *2016-03-16*
Published: *2016-08-15*

Stroben et al. BMC Medical Education (2016) 16:177
DOI 10.1186/s12909-016-0699-9

BMC Medical Education

RESEARCH ARTICLE Open Access

A simulated night shift in the emergency room increases students' self-efficacy independent of role taking over during simulation

Fabian Stroben[1,2*], Therese Schröder[1,2], Katja A. Dannenberg[1,3], Anke Thomas[2], Aristomenis Exadaktylos[4] and Wolf E. Hautz[4]

Abstract

Background: Junior doctors do not feel well prepared when they start into postgraduate training. High self-efficacy however is linked to better clinical performance and may thus improve patient care. What factors affect self-efficacy is currently unknown. We conducted a simulated night shift in an emergency room (ER) with final-year medical students to identify factors contributing to their self-efficacy and thus inform simulation training in the ER.

Methods: We simulated a night in the ER using best educational practice including multi-source feedback, simulated patients and vicarious learning with 30 participants. Students underwent 7 prototypic cases in groups of 5 in different roles (leader, member and observer). Feeling of preparedness was measured at baseline and 5 days after the event. After every case students recorded their confidence dependent of their role during simulation and evaluated the case.

Results: Thirty students participated, 18 (60 %) completed all surveys. At baseline students feel unconfident (Mean −0.34). Feeling of preparedness increases significantly at follow up (Mean 0.66, $p = 0.001$, $d = 1.86$). Confidence after simulation is independent of the role during simulation ($F(2,52) = 0.123$, $p = 0.884$). Observers in a simulation can estimate leader's confidence independent of their own ($r = 0.188$, $p = 0.32$) while team members cannot ($r = 0.61$, $p < 0.001$).

Conclusions: Simulation improves self-efficacy. The improvement of self-efficacy is independent of the role taken during simulation. As a consequence, groups can include observers as participants without impairing their increase in self-efficacy, providing a convenient way for educators to increase simulation efficiency. Different roles can furthermore be included into multi-source peer-feedback.

Keywords: Medical education, Undergraduate education, Simulation-based education, Emergency medicine, High-fidelity simulation, Self-assessment, Self-efficacy

Background

Junior doctors do not feel well prepared when they start into postgraduate training [1–3] independent of their objective performance [4]. Next to the accuracy of a diagnosis, adequate confidence in this diagnosis however is a necessity for safe and effective patient care. Too little confidence in an accurate diagnosis may harm patients through the delay of necessary treatment and unnecessary and potentially harmful additional investigations.

While the relationship between confidence and tendency to act applies to all of medicine, it is especially relevant to emergency medicine, where delayed action may have severe consequences.

Situational confidence (or self-efficacy) is a key factor to determine what actions one may take [5]. As an individual's reliance on personal abilities to succeed in a given challenge, self-efficacy increases the likelihood of that individual's actions actually occurring [6, 7]. By contrast, low

* Correspondence: fabian.stroben@charite.de
[1]Lernzentrum (Skills Lab), Charité-Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany
[2]Department of Gynecology and Obstretics CCM & CVK, Charité-Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany
Full list of author information is available at the end of the article

self-efficacy and resulting distress is argued to contribute to mental health problems [8, 9].

Several factors have been identified to influence self-reported feelings of preparedness. The percentage of graduates not feeling well prepared for clinical work differs strongly between countries [3, 10, 11] implying a great impact of educational systems and practices. Factors known to contribute to higher feelings of preparedness include frequent and immediate feedback [12], theoretical education counterbalanced with practice training, good skills education and training in diagnostic decision-making [10].

Simulation is a teaching format that may (and should) contain all four of those elements [13, 14] and thus should affect individual feelings of preparedness besides the well-known effects on objective performance [15]. Another teaching format known to increase self-efficacy includes observational or vicarious learning which is as effective as hands-on training in the acquisition of practical skills [16].

The aim of our study was to develop a best practice simulation session and evaluate the effect of simulation on the development of students' feelings of preparedness. We further aimed to identify factors within the simulation that affect confidence and feelings of preparedness in order to design a well-balanced simulation, budgeting both costs and educational effectiveness. To identify such factors we focused on the role students take over during the simulation and differences between self-reported confidence and confidence judged by peers.

Methods

Study design

A six-hour simulation session took place in 2013 at Charité – Universitätsmedizin Berlin as a night shift in a simulated emergency room (ER). The ER consisted of several rooms and an ambulance vehicle. Each room hosted a different simulated case of a total of seven. We invited students in their final year of medical school to participate.

Participants were randomized into teams of five. Each team rotated through each of the scenarios, thus seeing seven different patients, each presenting a typical ER case. Each group was staffed with a peer tutor who counseled on teamwork in between scenarios, helped with logistics and ensured participants filled in evaluations. Each room was staffed with a case tutor who ran the simulation scenario. Before starting each scenario, the group decided on a team leader, team members and observers. Feedback was given after each scenario. Figure 1 illustrates the study design.

At the beginning of the night, participants completed a questionnaire on possible confounder and self-reported feeling of preparedness in different medical specialties together with an informed consent form. Directly after each scenario, all active members recorded their confidence individually before feedback was given. Furthermore team members and observers estimated the team leader's confidence. After feedback, participants evaluated the quality of the given feedback and of the simulation overall. At the end of the night, participants filled in a final evaluation focusing on overall quality of the simulations. Participants were further asked to complete a second questionnaire of self-reported feelings of preparedness five days after the event in an online survey similar to the first questionnaire.

All evaluations (forms available upon request) were conducted using Likert scales for each item ranging from "totally agree" to "totally disagree". We coded the responses on a numerical scale ranging from +3 to −3 with 0 equating "neutral". All but the last evaluation forms were filled in on paper during the simulation, the final questionnaire "feeling of preparedness 2" was conducted online using lime survey software.

Participants

Medical students who had completed their fifth year of medical school (around 600 total) were invited to take part in the night shift. In Germany, five years of medical studies are conducted at university, the sixth and final year is spent in internships at different hospitals. Participants were chosen on a first come-first served basis through an online registration. 30 participants were randomized into six teams stratified by gender using a computer-generated randomization list. Participants were greeted in a general assembly and informed about the course of events of the night. After completing the written informed consent including participant information, information about opportunities to withdraw and possible consequences of withdrawl (none), they split into teams. The study was approved by the institutional office for data protection at Charité Berlin and deemed exempt from ethical review under local legislation, because it does not involve patients.

Cases

Cases were drafted following national and international guidelines and chosen based on learning objectives from a German national consensus statement [17, 18]. Each case represented a common ER patient. There were more diagnostic investigations available than necessary per case in order to ensure an uninterrupted simulation. A checklist was developed for each scenario to guide feedback by peer observers and case tutors.

All cases started with a presentation by the case tutor who enacted ER staff reporting a patient to the on-call physicians. Simulated Patients (SP) or simulators were placed as required by the scenario. Guideline-oriented therapy including airway-management was possible in all cases. I.v. medication and/or oxygen could be administered

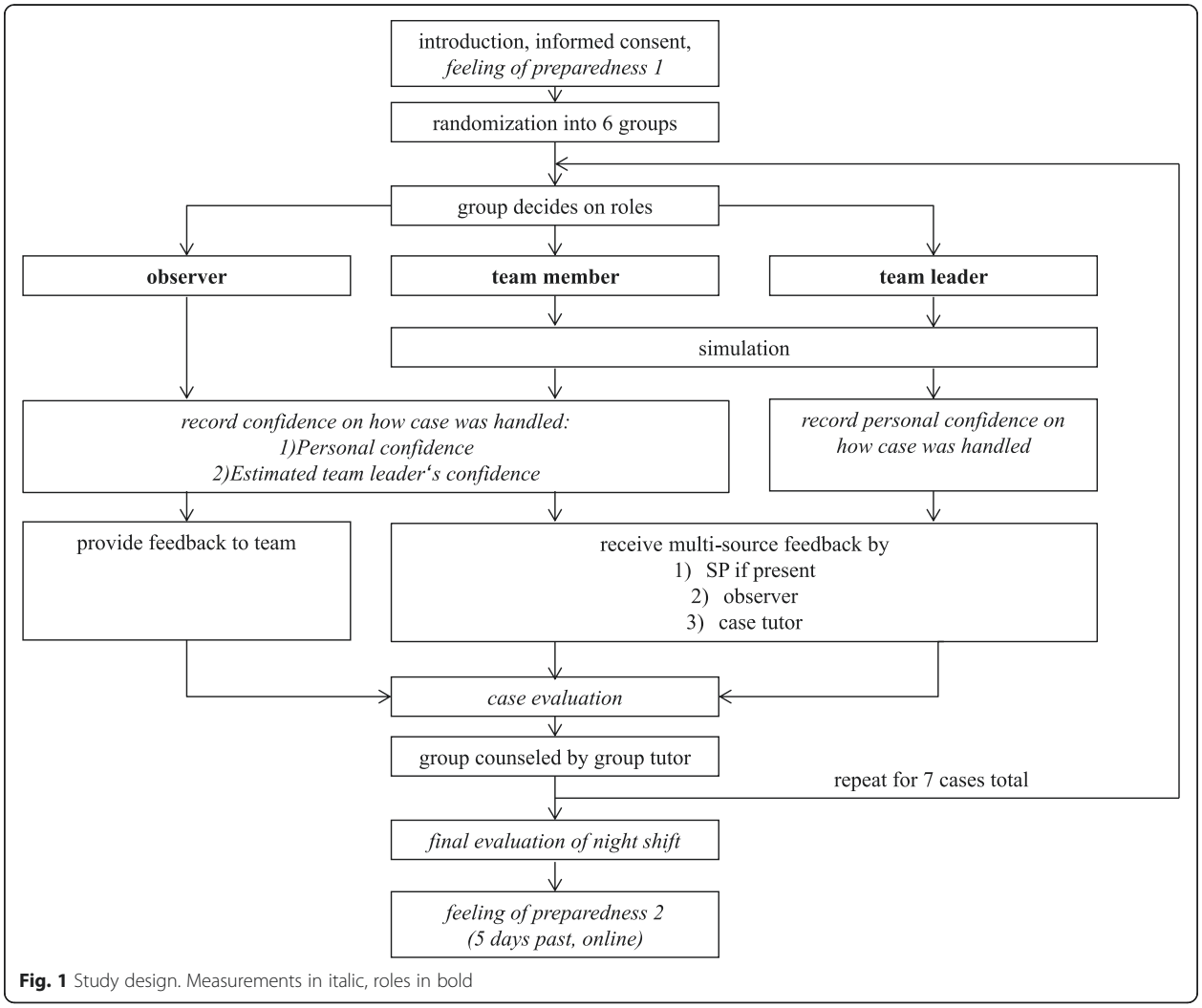

Fig. 1 Study design. Measurements in italic, roles in bold

if required. An overview of the developed cases is provided in Table 1.

Implementation

SPs were trained for five case scenarios (pulmonology, cardiology, neurology, urology and surgery 2, see Table 1). To guarantee an appropriate level of fidelity both hybrid simulations and mechanical simulators were used [14, 19]. To represent a real time course of events, all laboratory orders and radiology inquiries had to be requested by phone and/or in written form. The operational headquarters delayed their answer depending on the requested examination. Participants finished a scenario by arranging for the patient to be transferred to a ward or to be discharged. Each scenario lasted approximately 30–45 min including feedback. Additional technical details of the simulation, a detailed description of every case and used guidelines are provided as supporting information (see Additional files 1 and 2).

Roles of participants

For each scenario students took one of three roles:

- the *team leader* was responsible for the entire process – coordinating the team, choosing the right diagnosis and treating the patient accordingly.
- the *team member* was an active part of the group and supported the team leader throughout the process of finding the right diagnosis and treating the patient.
- the *team observer* observed the team using a checklist and provided feedback afterwards.

Roles within the group changed with each scenario so that at the end of the night each student had at least

Table 1 Cases and simulation settings

Discipline	Diagnosis (guidelines as sources)	Mode of simulation	Anticipated course of management
Pulmology	Exacerbated COPD	SP with examination possible	Chest X-Ray, blood-gas analysis, continous monitoring
Neurology	Ischemic media-stroke	SP with examination possible	CCT, continous monitoring
Cardiology	STEMI & non-sustained ventricular tachykardia	SP with examination possible	12-channel ECG, enzymes, continous monitoring
Anaesthesia	Ventricular fibrillation following STEMI	simulator-based approach	continous monitoring, ACLS
Surgery 1	Hemodynamic instable ruptured spleen	simulator-based approach with advanced monitoring	ATLS with FAST, continous monitoring
Urology	Urinary tract infection & pregnancy	SP with examination and sonography possible	urin test, ultrasound and gynaecological referal
Surgery 2	Head laceration	SP with examination and preparation of wound possible	Stitching of the wound

once taken on each role. Each group could freely develop their teamwork throughout the night shift. A peer tutor supervised and counseled the group.

Multi-source feedback

We used multi-sourced feedback [20] given by observers with specific assignments:

- the *SP* focused on communication using the Calgary-Cambridge Observation Guide (CCOG) to guide his or her feedback [21, 22].
- the *team observer* gave checklist-based feedback in order to provide the team with external observations but also to increase active monitoring of the simulation for his or her personal learning effect.
- the *case tutor* focused on the decision-making process with regards to medical content using case-specific checklists.
- the *peer tutor* focused on general teamwork and the development of team dynamics and gave feedback in a distinct setting to separate it from the case scenarios.

All tutors are trained in giving feedback and have extensive experience in peer teaching. Participants had experience giving and receiving feedback through curricular events. All SPs are trained regularly.

Statistical analysis

Collected data were analyzed with IBM SPSS Statistics 21.0 (SPSS Inc, Chicago, IL, USA). All data were first analyzed descriptively (mean, standard deviation). Confounder for feeling of preparedness were analyzed with Mann-Whitney-U-Tests, for differences between feeling of preparedness 1 and feeling of preparedness 2 we used a related sample Wilcoxon signed-rank test. For analysis of role and confidence we conducted a repeated measures analysis of variance (ANOVA), results of which we report as *F*- and *p*-values. Correlations between roles were analyzed with Pearson-correlations. Significance was defined as $p < 0.05$, Cohen's d was calculated as effect size. We further used G*Power, version 3.1.9.2 [23] to calculate the power achieved. We determined a gain of 0.51 on the Likert scale used from before to after the simulation as the smallest meaningful difference, because such a change would imply that participants chose one point better on the scale slightly more often than expected by chance. The primary dataset is provided as supporting information (see Additional file 3).

Results

A total of 30 students (20 female) participated in the simulation. Three participants had previous medical experience as a paramedic (2) or nurse (1). All 30 available places were booked up after 30 min in the online registration.

Feeling of preparedness

Participants feel rather ill prepared to care for patients before the simulation regardless of specialty (Mean −0.34) with no significant differences between gender ($p = 0.075$) or age ($p = 0.9$).

Right after each case students feel confident in their actions and with how they handled the case (Mean 0.95). All participants completed all surveys during the event (100 % response rate), 18 of the 30 participants (60 %) completed the online survey five days after the simulation and showed a significant increase in their general feeling of preparedness compared to before the simulation ($p = 0.001$). Participants now report to generally feel prepared (Mean 0.66); the effect is large ($d = 1.86$).

We analyzed these overall effects for every implemented discipline during simulation and found significant increases in the feeling of preparedness in anesthesiology, urology and taking history (see Table 2). The power of this study to detect a change in feeling of preparedness of 0.51 or greater was 99.79 %.

Table 2 Feeling of preparedness and change from before to five days after simulation

Discipline	Feeling of prepardness Baseline (Mean & SD)	Feeling of prepardness Follow Up (Mean & SD)	*p*-value
Overall	−0.34 (0.49)	0.66 (0.59)	0.001**
Taking History	1.27 (1.02)	1.72 (0.9)	0.035*
Anaesthesiology	0.14 (1.06)	1.17 (0.62)	<0.001***
Urology	−0.77 (1.25)	0.28 (1.53)	0.013*
Cardiology	−0.1 (1.06)	0.28 (1.13)	0.145
Pulmonology	−0.4 (0.97)	0.11 (1.13)	0.07
Surgery	0.13 (1.33)	0.83 (1.3)	0.101
Neurology	−0.47 (1.07)	0.22 (1.11)	0.1

Likert scales from +3 (totally agree) to −3 (totally disagree) we used for each item. *$p < 0,05$, **$p = 0,001$, ***$p < 0,001$

Role and confidence

In a repeated measures ANOVA with case as the within subject and role as the between subject factor, the self-reported confidence of participants is independent of their role during the simulation ($F(2,52) = 0.123$ $p = 0.884$). Both, team members and observers, are equally capable of judging the team leader's confidence independent of their own role ($F(2,52) = 2.055$ $p = 0.138$). How an active team member judges the team leader's confidence is in part predicted by his or her own confidence ($r = 0.61$; $p < 0.001$) while the confidence of the team leader judged by the passive observers does not correlate to the observer's personal confidence ($r = 0.188$; $p = 0.32$).

General evaluation

The simulation was evaluated very positively. Students were especially satisfied with how their peer tutors cared for them (Mean 2.93), how the SPs portrayed the patients, the difficulty of the scenarios and their opportunity to apply knowledge learned in medical school (all Mean > 2.7). The quality of the simulation was judged as very good (Mean 2.58). The ratings of each scenario right after the case correspond to the overall evaluation of the night shift.

Students reported to take most out of the feedback given by the case tutors (Mean 2.5) and slightly less out of the feedback by SPs and observing team members (both Mean 2.0).

Discussion

In line with previous findings [11, 24], especially in acute care [2], this study identifies a low feeling of preparedness among medical school graduates with results comparable to previous German [10] and British [3] studies. Our results provide evidence that even a relatively short simulation lasting just one night is effective in increasing students self-efficacy significantly as we observed an overall effect size of $d = 1.86$. Cohen himself suggested to classify effects as small when *Cohen's* $d > 0.2$, as medium when $d > 0.5$ and as large when $d > 0.8$ [25].

Intentionally including phases with observational tasks instead of active participation into the simulation may very well explain the simulations large increase simulation efficiency. Stegmann et al. previously demonstrated that hands-on-learning is as efficient as vicarious learning in the acquisition of complex manual skills [16] and Bloch and Bloch successfully used this method in ER-training sessions [26]. Active observation however is a requirement for vicarious learning [27] and giving feedback further enhances it [28]. Our results show that the effect of vicarious learning extends beyond knowledge and skills acquisition and affects situational confidence and ultimately the feeling of preparedness which we found to be unrelated to a learner's role during simulation. This provides a convenient opportunity for educators to increase group size in simulation with distributed, changing roles among participants and can influence the ratio of staff vs. participants to a more economic one. Furthermore, a recent study has demonstrated a large increase in diagnostic accuracy if patients are diagnosed by teams instead of individuals [29], further increasing the necessity to train medical staff in collaboration and to improve familiarity between ER-teammates, which was found to be surprisingly low in a recent observational study [30].

Training in the night may also be beneficial – nighttime hours are a neglected part of physicians training and may help to better prepare medical graduates for clinical settings [31] and reduce subjective stress of residents working on nighttime [24].

The observation that students significantly gained confidence in history taking may be explained by the facts that a) history taking was required in all cases presented during the night shift and students thus had ample opportunity to practice and b) history taking is directly observable to fellow students and tutors and participants may thus have received plenty of feedback regarding their interviewing skills. We can however only speculate as to why students feeling of preparedness improved for some (i.e. anesthesiology and urology) but not other (i.e. cardiology, pulmonology, surgery and neurology) disciplines and reasons might be discipline-specific. The change in urology may well be attributed to the fact that students hear little to nothing about this discipline during

their course of studies [17, 18], while the increased feeling of preparedness in anesthesiology may be due to the high prevalence of algorithms in this discipline. However, the factors that determine changes in the feeling of preparedness warrant further study.

Beyond their implication for simulation practice, our results may also effect future studies of physician confidence. The observation that a team leaders self-reported confidence is not significantly different from his or her confidence judged by observers indicates an equivalence of self-reported and behavioral indicators of situational confidence. This finding further justifies the use of both measures in research on situational confidence, elsewhere termed self monitoring [32]. How the previous experimental finding, that discrepancies in confidence between team members is predictive of team failure [30], translates to real-world medical practice is currently explored in different studies [33]. Although we also did not find differences between team leader's confidence and their confidence judged by team members, team members account of the leader's confidence correlates to their own and should thus not be regarded as a valid measure.

Limitations

Because of the high personal effort and costs per participant, only a small number of students were included into the night shift simulation and this pilot study. This might be one reason for non-significant changes in feeling of preparedness in some disciplines. Achieved power however was adequate, thus implying that increasing sample size would likely only lead to the identification of irrelevant findings.

Further, one could argue that the feeling of preparedness is not necessarily linked to objective performance [34], an aspect discussed controversially [35] since self-efficacy is known to become a self-fulfilling prophecy by actually raising the chances of success on a given task [36]. In line with this model of self-efficacy, Bloch [26] and Schubert [37] both found good performance to be associated with high levels of self-reported feelings of preparedness.

Conclusion

Best-practice simulation increases the feeling of preparedness in medical students but remains expensive in the conceptual process. Assigning participants to different roles during simulation is a convenient way to increase group size. These roles have no negative influence on the increase in self-efficacy and provide an opportunity for implementing multi-source peer-feedback. The feeling of preparedness of the active team members and leader also is apparent to observers and can be used as part of a debriefing after a simulation.

Additional files

Additional file 1: Case descriptions (used Guidelines indicated as references). (PDF 87 kb)

Additional file 2: Technical details. Diagnosis and technical details for every case. (PDF 70 kb)

Additional file 3: Dataset of the study. (XLS 65 kb)

Abbreviations

ANOVA, analysis of variance; CCOG, Calgary-Cambridge observation guide; ER, emergency room; SP, simulated patient

Acknowledgments

We want to acknowledge all involved tutors of the Lernzentrum, especially Katharina Ziegeler, Leopold Vorwerk and Malte Stuhlert for participation in conception and organisation of the study. We also want to thank the Simulated Patients Program of Charité especially Bea Kampel for her support in the use of standardized patients. Furthermore we would like to acknowledge Kerry Goulston, Sydney, and Sebastian Höfer and Miriam Rüsseler, both from Frankfurt, for their constructive critique on an earlier version of the manuscript.

Funding

The study was partly funded by an unrestricted grant from Deutsche Ärztefinanz to the Skills Lab at Charite. The sponsor did not interfere with the conception and conduction of the study, analysis of data or production of the manuscript.

Availability of data and materials

The dataset supporting the conclusions of this article is included within the article (and its additional files).

Authors contributions

All authors have read and approved the manuscript. FS, KAD and WEH were responsible for the study conception and design. KAD, FS and TS performed the data collection. WEH performed the data analysis. All authors contributed to data interpretation. FS, TS and KAD were responsible for the drafting of the manuscript. All authors made critical revisions to the manuscript for important intellectual content. WEH obtained funding. WEH and AT provided administrative, technical or material support. WEH, AT and AE supervised the study.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

The study was approved by the institutional office for data protection at Charité Berlin and deemed exempt from ethical review under local legislation, because it does not involve patients.

Author details

[1]Lernzentrum (Skills Lab), Charité-Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany. [2]Department of Gynecology and Obstretics CCM & CVK, Charité-Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany. [3]Department of Emergency Medicine at Campus Benjamin Franklin, Charité-Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany. [4]Universitäres Notfallzentrum, Inselspital Bern, 3010 Bern, Switzerland.

Received: 15 April 2016 Accepted: 16 June 2016
Published online: 15 July 2016

References

1. Morgan PJ, Cleave-Hogg D. Comparison between medical students' experience, confidence and competence. Med Educ. 2002;36:534–9.
2. Tallentire VR, Smith SE, Skinner J, Cameron HS. The preparedness of UK graduates in acute care: a systematic literature review. Postgrad Med J. 2012;88:365–71.

3. Goldacre MJ, Taylor K, Lambert TW. Views of junior doctors about whether their medical school prepared them well for work: questionnaire surveys. BMC Med Educ. 2010;10:78.
4. Tokuda Y, Goto E, Otaki J, Jacobs J, Omata F, Obara H, Shapiro M, Soejima K, Ishida Y, Ohde S, Takahashi O, Fukui T. Undergraduate educational environment, perceived preparedness for postgraduate clinical training, and pass rate on the National Medical Licensure Examination in Japan. BMC Med Educ. 2010;10:35.
5. Woods JL, Pasold TL, Boateng BA, Hensel DJ. Adolescent health care and the trainee: roles of self-efficacy, standardized patients, and an adolescent medicine rotation. Simul Healthc. 2013;8:359–67.
6. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol. 2001;52:1–26.
7. Conner M, Norman P: Predicting Health Behaviour. Volume 24. Berkshire, England: Open University Press; 2005.
8. Tyssen R, Vaglum P. Mental health problems among young doctors: an updated review of prospective studies. Harv Rev Psychiatry. 2002;10:154–65.
9. Hsu K, Marshall V. Prevalence of Depression and Distress in a Large Sample of Canadian Residents, Interns, and Fellow. Am J Psychiatry. 1987;144:1561–6.
10. Ochsmann EB, Zier U, Drexler H, Schmid K. Well prepared for work? Junior doctors' self-assessment after medical education. BMC Med Educ. 2011;11:99.
11. Cave J, Woolf K, Jones A, Dacre J. Easing the transition from student to doctor: how can medical schools help prepare their graduates for starting work? Med Teach. 2009;31:403–8.
12. Hattie J, Timperley H. The Power of Feedback. Rev Educ Res. 2007;77:81–112.
13. Cook DA, Hamstra SJ, Brydges R, Zendejas B, Szostek JH, Wang AT, Erwin PJ, Hatala R. Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. Med Teach. 2013;35:e867–98.
14. Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach. 2005;27:10–28.
15. Cortegiani A, Russotto V, Montalto F, Iozzo P, Palmeri C, Raineri SM, Giarratano A. Effect of High-Fidelity Simulation on Medical Students' Knowledge about Advanced Life Support: A Randomized Study. PLoS One. 2015;10:e0125685.
16. Stegmann K, Pilz F, Siebeck M, Fischer F. Vicarious learning during simulations: is it more effective than hands-on training? Med Educ. 2012;46:1001–8.
17. Schnabel KP, Boldt PD, Breuer G, Fichtner A, Karsten G, Kujumdshiev S, Schmidts M, Stosch C. A consensus statement on practical skills in medical school - a position paper by the GMA Committee on Practical Skills. GMS Z Med Ausbild. 2011;28:Doc58.
18. Blaum WE, Dannenberg KA, Friedrich T, Jarczewski A, Reinsch AK, Ahlers O. The practial use of the consensus statement on practical skills in medical school–a validation study. GMS Z Med Ausbild. 2012;29:Doc58.
19. May W, Park J, Lee J. A ten-year review of the literature on the use of standardized patients in teaching and learning: 1996–2005. Med Teach. 2009;31:487–92.
20. Wood L, Hassell A, Whitehouse A, Bullock A, Wall D. A literature review of multi-source feedback systems within and without health services, leading to 10 tips for their successful design. Med Teach. 2006;28:e185–91.
21. Kurtz SM, Silverman JD. The Calgary-Cambridge Referenced Observation Guides: an aid to defining the curriculum and organizing the teaching in communication training programmes. Med Educ. 1996;30:83–9.
22. Kurtz S, Silverman J, Benson J, Draper J. Marrying Content and Process in Clinical Method Teaching: Enhancing the Calgary–Cambridge Guides. Acad Med. 2003;78:802–9.
23. Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007;39:175–91.
24. Brennan N, Corrigan O, Allard J, Archer J, Barnes R, Bleakley A, Collett T, de Bere SR. The transition from medical student to junior doctor: today's experiences of Tomorrow's Doctors. Med Educ. 2010;44:449–58.
25. Cohen J. A power primer. Psychol Bull. 1992;112:155–9.
26. Bloch SA, Bloch AJ. Simulation training based on observation with minimal participation improves paediatric emergency medicine knowledge, skills and confidence. Emerg Med J. 2015;32:195–202.
27. Chi MTH, Roy M, Hausmann RGM. Observing tutorial dialogues collaboratively: insights about human tutoring effectiveness from vicarious learning. Cogn Sci. 2008;32:301–41.
28. Grierson LEM, Barry M, Kapralos B, Carnahan H, Dubrowski A. The role of collaborative interactivity in the observational practice of clinical skills. Med Educ. 2012;46:409–16.
29. Hautz WE, Kämmer JE, Schauber S, Spieß CD, Gassmeier W. Diagnostic Performance by Medical Students Working Individually or in Teams. JAMA. 2015;313:303–4.
30. Patterson PD, Pfeiffer AJ, Lave JR, Weaver MD, Abebe K, Krackhardt D, Arnold RM, Yealy DM. How familiar are clinician teammates in the emergency department? Emerg Med J. 2015;32:258–62.
31. Hanson JT, Pierce RG, Dhaliwal G. The New Education Frontier: Clinical Teaching at Night. Acad Med. 2013;89:1–4.
32. Eva K, Regehr G. "I'll never play professional football" and other fallacies of self-assessment. J Contin Educ Health Prof. 2008;28:14–9.
33. Hautz SC, Schuler L, Kämmer JE, Schauber SK, Ricklin ME, Sauter TC, Maier V, Birrenbach T Exadaktylos A, Hautz WE. Factors predicting a change in diagnosis in patients hospitalised through the emergency room: a prospective observational study. BMJ Open. 2016;6:e011585.
34. Barnsley L, Lyon PM, Ralston SJ, Hibbert EJ, Cunningham I, Gordon FC, Field MJ. Clinical skills in junior medical officers: a comparison of self-reported confidence and observed competence. Med Educ. 2004;38:358–67.
35. Duns G, Weiland T, Crotty B, Jolly B, Cuddihy H, Dent A. Self-rated preparedness of Australian prevocational hospital doctors for emergencies. Emerg Med Australas. 2008;20:144–8.
36. Eva K, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. Acad Med. 2005;80(10 Suppl):S46–54.
37. Schubert A, Tetzlaff JE, Tan M, Ryckmann V, Mascha E. Consistency, Inter-rater Reliability, and Validity of 441 Consecutive Mock Oral Examinations in Anesthesiology. Anesthesiology. 1999;91:288–98.

c. Shame in Medical Education: A Randomized Study of the Acquisition of Intimate Examination Skills and Its Effect on Subsequent Performance

Wolf E. Hautz, Therese Schröder, Katja A. Dannenberg, Maren März, Henrike Hölzer, Olaf Ahlers & Anke Thomas (2017) Shame in Medical Education: A Randomized Study of the Acquisition of Intimate Examination Skills and Its Effect on Subsequent Performance, Teaching and Learning in Medicine, 29:2, 196-206, DOI: 10.1080/10401334.2016.1254636

5. Lebenslauf

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

6. Komplette Publikationsliste

Artikel

The future of practical skills in undergraduate medical education - an explorative Delphi-Study.
Dannenberg KA, Stroben F, Schröder T, Thomas A, Hautz WE.
GMS J Med Educ. 2016 Aug 15;33(4):Doc62. doi: 10.3205/zma001061.
PMID: 27579362

A simulated night shift in the emergency room increases students' self-efficacy independent of role taking over during simulation.
Stroben F, Schröder T, Dannenberg KA, Thomas A, Exadaktylos A, Hautz WE.
BMC Med Educ. 2016 Jul 15;16:177. doi: 10.1186/s12909-016-0699-9.
PMID: 27421905

Shame in Medical Education: A Randomized Study of the Acquisition of Intimate Examination Skills and Its Effect on Subsequent Performance.
Hautz WE*, Schröder T*, Dannenberg KA, März M, Hölzer H, Ahlers O, Thomas A.
Teach Learn Med. 2017 Jan 4:1-11. doi: 10.1080/10401334.2016.1254636.
PMID: 28051893
*geteilte Erstautorenschaft

Poster

3. Berliner Posterkongress „Wissenschaftliches Arbeiten im Reformstudiengang Medizin“, Charité - Universitätsmedizin Berlin: „Zusammenhang zwischen Scham und manueller Performance bei der Brustuntersuchung“ – präsentiert von Therese Schröder, 07/2012

40th Annual Conference Association for Medical Education in Europe (AMEE), Prag: „How does shame affect acquisition of manual skills?“ – präsentiert von Wolf Blaum, 08/2013

Präsentation

Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA), Graz, 26.-28.09.2013. Düsseldorf: German Medical. Science GMS Publishing House; 2013. DocV13_01. DOI: 10.3205/13gma223, URN: urn:nbn:de:0183-13gma2231; „Scham, Stress und der Erwerb praktischer Fertigkeiten“ – präsentiert von Wolf Blaum, 09/2013

7. Danksagung

Ich möchte mich bei all denen bedanken, die zum Gelingen dieser Dissertation beigetragen haben, insbesondere bei den Teilnehmern der drei Studien, ohne die es diese Arbeit nicht geben würde.
Ich danke Prof. Dr. med. Anke Thomas für die Möglichkeit zu dieser Promotion, die fachliche Unterstützung und die finale Korrektur der Dissertationsarbeit.
Mein besonderer Dank gilt Wolf Hautz für die fabelhafte Betreuung über viele Jahre, die Einführung in das wissenschaftliche Arbeiten und die Begeisterung für die medizinische Lehre. Unabhängig davon, wo wir beide uns jeweils gerade befanden, konnte ich mich immer auf seine prompten und hilfreichen Antworten verlassen!
Des weiteren danke ich meinen Co-Autoren für die konstruktive Zusammenarbeit: Aristomenis Exadaktylos, Fabian Stroben, Henrike Hölzer, Katja Dannenberg, Maren März, Olaf Ahlers.
Zudem bedanke ich mich für die Unterstützung im Großen wie im Kleinen bei Agnes Waldschütz, Andrea Antolic, Anne-Kathrin Reinsch, Antonia Janssen, Elisabeth von Tucher, Friederike Weber, Henriette Kiep, Janko Settler, Johanna Ramm, Larry Hinkson, Mary McKee, Nora Berger, Silvia Fisch, Stefan Schauber, Susanne Werner.
Meiner Familie und meinen Freunden danke ich für die geduldige Unterstützung, ihre offenen Ohren (und Haustüren!) und motivierenden Worte über viele Jahre.