

Aus dem Zentrum für Operative Medizin der Universität zu Köln
Klinik und Poliklinik für Allgemein-, Viszeral- und Tumorchirurgie
und Transplantationschirurgie
Direktorin: Universitätsprofessorin Dr. med. C. Bruns

Der Effekt von Ablenkungen durch Telefonanrufe auf die Leistungsfähigkeit von Medizinstudierenden während einer OSCE-Prüfung

Inaugural-Dissertation zur Erlangung der Doktorwürde
der Medizinischen Fakultät
der Universität zu Köln

vorgelegt von
Louisa Roentgen
aus Aachen

promoviert am 13.05.2026

Dekanin/Dekan: Universitätsprofessor Dr. med. G. R. Fink

1. Gutachter: Privatdozent Dr. med. R. K. Kleinert
2. Gutachter: Universitätsprofessor Dr. sc. hum. J. Koenig

Erklärung

Ich erkläre hiermit, dass ich die vorliegende Dissertationsschrift ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskriptes habe ich keine Unterstützungsleistungen erhalten.

Weitere Personen waren an der Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe einer Promotionsberaterin/eines Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertationsschrift stehen.

Die Dissertationsschrift wurde von mir bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die dieser Arbeit zugrunde liegende Studie ist durch die Arbeitsgemeinschaft Lehrforschung geplant und umgesetzt worden. An der Planung waren außer mir Dr. Rabi Datta, Dr. Robert Kleinert, Dr. Justus Toader, Dr. Thomas Dratsch sowie die Doktorand*innen Caroline Kroll, Martina Graupner, Moritz Zeeh und Nadja Jakovljevic beteiligt.

Die Datenerhebung als Prüfer*innen und die Beteiligung als Schauspieler*innen und Anrufer*innen erfolgten durch mich sowie Caroline Kroll, Martina Graupner, Moritz Zeeh, Nadja Jakovljevic und Thomas Dratsch

In der offiziellen OSCEII Prüfungsstation wurden außerdem professionelle Schauspieler*innen, welche von der Universität beauftragt und bezahlt wurden, eingesetzt.

Die Hauptverantwortung in der Planung und Koordination der Studienvorbereitung und Studiendurchführung lag bei mir.

Die statistische Auswertung erfolgte mit der Statistiksoftware Statistical Package for the Social Sciences (SPSS, Version 25; IBM, 2017) in Zusammenarbeit mit Dr. Thomas Dratsch. Es erfolgte eine umfassende Auswertung der Daten sowohl für die vorliegende Arbeit, als auch für einen vorab veröffentlichten Artikel. Ein Teil der Studiendaten wurde vorab veröffentlicht in dem Artikel „Effect of phone call distraction on the performance of medical students in an OSCE“¹.

In der vorliegenden Arbeit wurde das Gendersternchen verwendet unter Berücksichtigung der Empfehlungen der Universität zu Köln und der Medizinischen Fakultät. Das Gendersternchen wird vom Rektorat sowie im Leitfaden der Gleichstellungsbeauftragten der Universität zu Köln empfohlen, auch wenn dies noch nicht in das Regelwerk der deutschen Rechtschreibung aufgenommen wurde^{2,3}.

Erklärung zur guten wissenschaftlichen Praxis:

Ich erkläre hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten (Amtliche Mitteilung der Universität zu Köln AM 132/2020) der Universität zu Köln gelesen habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen.

Köln, den 30.11.2025

Danksagung

Zunächst möchte ich meinem Doktorvater Priv.-Doz. Dr. Robert Kleinert und meinem Betreuer Priv.-Doz. Dr. Rabi Datta danken, die mir ermöglicht haben diese Arbeit durchzuführen und mich in der Umsetzung betreut haben.

Auch anderen Mitgliedern der Arbeitsgemeinschaft Lehrforschung möchte ich für die Unterstützung insbesondere bei der Planung und Durchführung der Studie danken, hierzu gehören Dr. Justus Toader, Dr. Thomas Dratsch und Priv.-Doz. Dr. Patrick S. Plum.

Dr. med. Jan Griesinger hat mich ebenso in der Planungsphase unterstützt.

Für die Teamzusammenarbeit und gegenseitige Unterstützung möchte ich den anderen Doktorand*innen der Arbeitsgruppe Caroline Kroll, Martina Graupner, Moritz Zeeh und Nadja Jakovljevic danken.

Des Weiteren möchte ich mich bei den Mitarbeiter*innen des KISS - Kölner Interprofessionelles Skills Lab und Simulationszentrum bedanken. Mein Dank gilt besonders Sabine Bornemann und Christian Thrien, die eine Studienstation in OSCE I und OSCE II ermöglicht haben und mir jederzeit für organisatorische Fragen zur Seite standen.

Ich möchte außerdem meiner Familie, meinen Studienfreund*innen und meinem Partner danken, die immer für mich da waren.

Inhaltsverzeichnis

ABKÜRZUNGSVERZEICHNIS	9
1. ZUSAMMENFASSUNG	10
2. EINLEITUNG	11
2.1 Unterbrechungen, Ablenkungen und Multitasking im ärztlichen Alltag – Häufigkeit und Definitionen	11
2.2 Unterbrechungen von ärztlichen Tätigkeiten im klinischen Alltag durch technische Kommunikationsmittel	11
2.3 Allgemeiner Forschungsstand zu Auswirkungen von Ablenkungen durch Telefonanrufe im Alltag am Beispiel von Autofahren	13
2.4 Forschungsstand zu Auswirkungen von Ablenkungen und Unterbrechungen durch Telefonanrufe im Krankenhaus	14
2.5 OSCE-Prüfungen in der medizinischen Lehre	16
2.6 Fragestellungen und Ziel der Arbeit	18
3. MATERIAL UND METHODEN	19
3.1 Studiendesign	19
3.1.1. Kollektiv	19
3.1.2. Time series, within-group Design	20
3.1.3. Experimentelles Vorgehen	21
3.1.4. Aufgabenstellung	23
3.1.5. Zeitlicher Ablauf	24
3.1.6. Inhalt der Telefonanrufe	26
3.2 Dokumentation und Bewertung	27
3.3 Materialien	30
3.3.1. OSCE I	30
3.3.2. OSCE II	31
3.4 Aufgaben der beteiligten Doktorand*innen	32
3.5 Statistische Auswertung	32

4. ERGEBNISSE	33
4.1 Alter und Geschlecht der Proband*innen	33
4.1.1. OSCE I	33
4.1.2. OSCE II	33
4.2 Effekt der Telefonanrufe	33
4.2.1. Vergleich der erreichten Punktzahl, OSCE I	33
4.2.2. Vergleich der erreichten Punktzahl, OSCE II	33
4.2.3. Bearbeitung der Aufgabe im vorgegebenen Zeitrahmen	34
4.2.4. Unterschiede in der erreichten Punktezahl aufgeschlüsselt nach Items in der Checkliste	36
4.2.4.1 OSCE I	36
4.2.4.2 OSCE II	37
4.2.5. Korrelation zwischen der Leistung in der regulären Station und der Studienstation	39
4.3 Korrektes Wiedergeben der Informationen aus dem Telefonanruf	39
4.3.1. Unterschiede zwischen den Patientennamen	40
4.4 Prüfer*innen	41
4.4.1. Unterschiede zwischen den Prüfer*innen der Studie	41
4.4.2. Korrelation zwischen den Prüfer*innen der Studie und den Prüfer*innen der Universität	42
4.5 Ausbildung im medizinischen Bereich	43
4.5.1. OSCE I	43
4.5.2. OSCE II	43
4.6 Muttersprache der Studierenden	44
4.6.1. OSCE I	44
4.6.2. OSCE II	44
4.7 Regelmäßige sportliche Betätigung	45
4.8 Geschwister	45
4.9 Videospiele	46
4.10 OSCE II Szenarien	46

5.	DISKUSSION	50
5.1	Zusammenfassung der Ergebnisse	50
5.2	Diskussion der Ergebnisse: Effekt der Telefonanrufe auf die Leistung der Proband*innen und eine Einordnung in den aktuellen Forschungsstand	50
5.2.1.	Effekt der Telefonanrufe auf die Bearbeitungsdauer der unterbrochenen Tätigkeit	52
5.2.2.	Effekt der Telefonanrufe auf Teilaufgaben der Prüfung und einzelne Arbeitsschritte	54
5.2.3.	Zeitpunkt der Unterbrechung als relevanter Faktor für den Effekt der Unterbrechung	55
5.2.4.	Effekt der Telefonanrufe auf die zweite Tätigkeit und auf Gedächtnisleistungen: Korrektes Wiedergeben der Informationen aus den Telefonanrufen	56
5.2.5.	Effekt der Telefonanrufe abhängig von Ausbildung und Berufserfahrung: Erfahrungen im Umgang mit Unterbrechung als relevanter Faktor für den Effekt von Unterbrechungen	59
5.2.6.	Effekt der Telefonanrufe abhängig von der Komplexität der primären und sekundären, unterbrechenden Tätigkeiten (Unterschiede zwischen den OSCE II Szenarien)	61
5.2.7.	Effekt der Telefonanrufe abhängig von der Variable Muttersprache	63
5.2.8.	Effekt der Telefonanrufe abhängig von den Variablen Sport, Geschwister, Videospiele	64
5.3	Konsequenzen von Unterbrechungen durch Telefonanrufe für die Patient*innensicherheit und ärztliche Tätigkeit	65
5.4	Lösungsansatz: Sensibilisierung von medizinischem Personal und Training im Umgang mit Unterbrechungen und Ablenkungen	66
5.5	Lösungsansatz: Human Factor und systemische Veränderungen	69
5.6	Diskussion der Methodik und Limitationen	71
5.6.1.	Studiendesign und interne Validität	71
5.6.2.	Bedrohungen für die interne Validität: Übertragungseffekte, Übungseffekte, Sensibilisierung	72
5.6.3.	Bedrohung für die interne Validität: Instrumentation	72
5.6.4.	Kritische Reflektion der Methodik in Bezug auf die Szenarien in OSCE II	74
5.6.5.	Externe Validität	76
5.6.6.	Übertragbarkeit der Ergebnisse der OSCE-Prüfung auf den ärztlichen Alltag	76
5.6.7.	Übertragbarkeit der Leistung der Proband*innen (Studierende) auf die Leistung von Ärzt*innen	77
5.7	Fazit	78
6.	LITERATURVERZEICHNIS	79

7.	ANHANG	93
7.1	Abbildungsverzeichnis	93
7.2	Tabellenverzeichnis	93
7.3	Materialien	94
7.3.1.	Fragebogen	94
7.3.2.	Materialien OSCE I	95
7.3.2.1.	Fallvignette OSCE I reguläre Station	95
7.3.2.2.	Fallvignette OSCE I Studienstation	96
7.3.2.3.	Aufteilung Blutzuckerwerte und Patientennamen	97
7.3.3.	Materialien OSCE II	98
7.3.3.1.	Fallvignette OSCE II reguläre Station	98
7.3.3.2.	Fallvignette OSCE II Studienstation	99
7.3.3.3.	Aufteilung Blutzuckerwerte und Patientennamen	100
7.3.3.4.	Szenarien Kombination	101
8.	VORABVERÖFFENTLICHUNGEN VON ERGEBNISSEN	102

Abkürzungsverzeichnis

CT	Computertomographie
DECT	Digital Enhanced Cordless Telecommunications
M	Mittelwert
OP	Operation
OSCE	Objective structured clinical examination
SD	Standardabweichung (Standard deviation)

1. Zusammenfassung

Ablenkungen und Unterbrechungen durch Telefonanrufe sind alltäglicher Bestandteil der Arbeit von Ärzt*innen in Kliniken ⁴⁻⁶. In nicht medizinischen Studien konnten negative Effekte durch Telefonanrufe besonders beim Autofahren nachgewiesen werden ^{7,8}. Bisherige Studien im medizinischen Bereich weisen ebenfalls auf potenziell negative Effekte durch Ablenkungen und Unterbrechungen hin ⁹⁻²⁴. Allerdings wurden Telefonanrufe als Quelle der Unterbrechungen und Ablenkungen in experimentellen Studien bisher nur bei simulierten Operationen untersucht ²⁵⁻²⁸.

Daher war das Ziel dieser Studie, die Effekte von Telefonanrufen auf alltägliche ärztliche Tätigkeiten in einer experimentellen Studie zu untersuchen sowie mögliche Konsequenzen für die Patientensicherheit und Handlungsempfehlungen für medizinisches Personal und die medizinische Lehre abzuleiten.

Zu diesem Zweck wurden im Rahmen der OSCE I und OSCE II Prüfungen an der medizinischen Fakultät in Köln Medizinstudierende in einer Studienstation durch zwei Telefonanrufe abgelenkt. In der Station mussten die Proband*innen einen venösen Zugang legen oder eine Anamnese und körperliche Untersuchung bei einem Schauspielpatienten/ einer Schauspielpatientin durchführen. Die Proband*innen mussten die Station einmal mit Unterbrechungen durch Telefonanrufe und einmal ohne Störungen durchlaufen. Die erreichte Punktzahl aus den beiden Stationen wurde miteinander verglichen.

Hierbei erzielten die Proband*innen sowohl in OSCE I, als auch in OSCE II in der Studienstation mit Telefonanrufen signifikant weniger Punkte.

Dieses Ergebnis ließ den Schluss zu, dass Ablenkungen und Unterbrechungen durch Telefonanrufe einen potenziell negativen Effekt auf die unterbrochene Tätigkeit und damit auch auf die Patient*innensicherheit haben.

Als Konsequenz sollte eine Sensibilisierung für mögliche Folgen von Unterbrechungen und Ablenkungen erfolgen und ein Training im Umgang mit diesen Ablenkungen ein Teil der medizinischen Weiterbildung werden. Zusätzlich kann ein systemischer Ansatz verfolgt werden, in dem das Arbeitsumfeld so angepasst wird, dass eine Reduktion der Anzahl von vermeidbaren Unterbrechungen durch Telefonanrufe erfolgt. Eine weitere Evaluation möglicher Einflussfaktoren auf den Effekt von Unterbrechungen und Ablenkungen durch Telefonanrufe und eine Evaluation verschiedener Maßnahmen sollte in Zukunft erfolgen.

2. Einleitung

2.1 Unterbrechungen, Ablenkungen und Multitasking im ärztlichen Alltag – Häufigkeit und Definitionen

Ärzt*innen sind im klinischen Alltag mit einer Vielzahl von Ablenkungen und Unterbrechungen konfrontiert ^{4,29}. So werden Internist*innen und Chirurg*innen bei ihrer Arbeit zwischen 2,9 und 8 mal pro Stunde unterbrochen und betreiben in 17- 20% der Arbeitszeit Multitasking ^{4,5,29–32}. Zum weiteren Verständnis folgt zunächst eine Einordnung der in der Literatur genutzten Definitionen und der in der vorliegenden Arbeit genutzten Begriffe.

Es wird zwischen zwei Formen von Multitasking unterschieden: serielles Multitasking bezeichnet die Durchführung von mehreren Aufgaben kurz hintereinander ^{33–35}. Diese Form des Multitasking wird auch als Task-Switching bezeichnet ^{33–35}. Dies entspricht dem Wechsel zwischen verschiedenen Aufgaben ^{33–35}. Gleichzeitiges Multitasking bezeichnet die parallele/simultane Ausführung von mehreren Tätigkeiten ^{33–35}. Dies wird auch als Dual-Task bezeichnet ^{33–35}. Single-Task bedeutet, dass nur eine einzelne Tätigkeit ausgeführt wird ^{33–35}. Die Begriffe Multitasking, Task-switching und Dual-task werden in der vorliegenden Arbeit wie im englischen Original verwendet.

In der englischsprachigen Literatur wird zudem zwischen Interruption und Distraction unterschieden. Eine häufig zitierte Definition von Healey et al. beschreibt Distraction als eine Unterbrechung der Aufmerksamkeit, sodass diese von einer aktuell ausgeführten Tätigkeit zu einem anderen Reiz oder einer anderen Tätigkeit gelenkt wird ^{36,37}. Bei einer Interruption wird eine zuvor ausgeführte Tätigkeit unterbrochen, um sich mit einem anderen Reiz oder einer anderen Tätigkeit auseinander zu setzen ^{36,37}. Der Begriff Distraction wird in dieser Arbeit mit Ablenkung übersetzt und Interruption mit Unterbrechung.

2.2 Unterbrechungen von ärztlichen Tätigkeiten im klinischen Alltag durch technische Kommunikationsmittel

Wie in 2.1 bereits aufgeführt sind Unterbrechungen, Ablenkungen und Multitasking alltäglicher Bestandteil der Arbeit von Ärzt*innen ^{4,5,29–32}. Neben persönlichen (face-to-face) Kontakten zählen Telefonanrufe und Pager zu einer der häufigsten Quellen von Unterbrechungen und Ablenkungen von medizinischem Personal ^{4–6}, sowohl während der Stationsarbeit ^{4,5} als auch im Operationssaal ^{17,37–39} oder in der Notaufnahme ^{4,9,12,40–43}.

In einer Studie von 1997 zur innerklinischen Kommunikationskultur in einem britischen Krankenhaus wurde festgestellt, dass in dem Beobachtungszeitraum der Studie durchschnittlich alle 18,5 Minuten eine Unterbrechung eines Arztes/ einer Ärztin durch eine Kontaktaufnahme per Stationstelefon oder Pager auftrat ⁴⁴.

Inzwischen gibt es eine Vielzahl an Kommunikationsmitteln, die in Kliniken eingesetzt werden und potenziell zu Unterbrechungen führen können. Dazu gehören neben alphanummerischen Pagern, mobile DECT-Diensttelefone und Smartphones^{45,46}. Insbesondere die Nutzung von Smartphones in Kliniken hat beim medizinischen Personal in den vergangenen Jahren weltweit zugenommen^{47–60}. Viele nutzen inzwischen auch ihr privates Smartphone bei ihrer klinischen Tätigkeit⁴⁷. Wie eine Studie in Großbritannien zeigte, nutzten 2015 92% der befragten Ärzt*innen ihr persönliches Smartphone für arbeitsbezogene Anrufe. 88% nutzen es, um arbeitsbezogene Nachrichten zu schreiben⁵⁹. Zudem werden Smartphones in Kliniken sowohl zur internen Kommunikation innerhalb des medizinischen Personals genutzt als auch zu Recherchezwecken. Außerdem kommt es im Arbeitsalltag zur Anwendung medizinischer Apps^{52,57,59–61}.

Eine ständige Erreichbarkeit scheint in der Krankenhausroutine sinnvoll zu sein, doch stellt sich gleichzeitig die Frage, inwieweit sich derartige Unterbrechungen und Ablenkungen auf die Konzentrations- und Leistungsfähigkeit des medizinischen Personals und damit auf die Patientensicherheit auswirken.

Unterbrechungen und Ablenkungen können, wie aus psychologischen Studien bekannt, zu Fehlern, einer längeren Bearbeitungsdauer von Aufgaben, einer höheren Arbeitsbelastung und zu Stress führen^{62–64}.

Der im Jahr 1999 veröffentlichte Bericht „To Err Is Human: Building a Safer Health System“ vom Committee on Quality of Health Care in America erregte mit seiner Forderung, die Sicherheit im Gesundheitssystem zu erhöhen großes Aufsehen⁶⁵.

In dem Bericht werden Unterbrechungen und Ablenkungen als eine von vielen möglichen Ursachen für Fehler genannt⁶⁶. Dies führte in den folgenden Jahren zu einer verstärkten Auseinandersetzung der Forschung mit dem Thema Patient*innensicherheit⁶⁵.

Zusammenfassend lässt sich feststellen, dass Kommunikationsmittel, insbesondere durch die Etablierung von Smartphones, zunehmend im klinischen Alltag genutzt werden.

Es ist denkbar, dass hierdurch die Häufigkeit von Unterbrechungen und Ablenkungen im klinischen Alltag durch Telefonanrufe noch weiter zunehmen wird.

Daher kommt der wissenschaftlichen Auseinandersetzung mit den potenziellen Folgen von Ablenkungen eine zunehmend große Bedeutung zu.

2.3 Allgemeiner Forschungsstand zu Auswirkungen von Ablenkungen durch Telefonanrufe im Alltag am Beispiel von Autofahren

Inwiefern sich der Gebrauch von Mobiltelefonen im Alltag auswirkt, wird in Studien zum Einfluss von Mobiltelefon-Gebrauch beim Autofahren besonders deutlich. Dies wurde in vielfachen Studien untersucht ^{7,8}. Mehrere Metaanalysen zeigten, dass sich Ablenkungen durch Telefongespräche beim Autofahren negativ auf das Fahrverhalten der Proband*innen der untersuchten Studien auswirkten ^{7,8}. Vor allem die Reaktionszeit war im Vergleich zum Autofahren ohne gleichzeitiges Telefongespräch verlangsamt ^{7,8}. Es kam häufiger zu Kollisionen⁷. In experimentellen Studien konnten weniger Detektionsaufgaben erfolgreich bearbeitet werden ⁷. Diese negativen Effekte wurden sowohl bei Gesprächen festgestellt, bei denen der Fahrer/ die Fahrerin das Telefon in der Hand hielt, als auch bei Gesprächen über Freisprechanlagen ^{7,8,67,68}.

Ergänzend wurden in neurophysiologischen Studien Magnetenzephalographien und funktionelle Magnetresonanztomographien eingesetzt, um zu beurteilen, wie sich die Hirnaktivität verändert, wenn Proband*innen simuliert Auto fahren und gleichzeitig eine zweite Aufgabe (meist das Beantworten von Fragen über ein Head-Set) bearbeiten. Es konnte ein signifikanter Wechsel der aktivierten Gehirnregionen, unter anderem von occipitalen zu frontoparietalen Regionen, beobachtet werden ⁶⁹.

Sanbonmatsu et al. zeigten in ihrer Studie ebenfalls, dass während einer simulierten Autofahrt und einem gleichzeitigen Telefongespräch mehr Fehler gemacht wurden ⁷⁰. Nach der Fahrt wurden die selbsteingeschätzte Fahrsicherheit und die Anzahl der Fehler der Proband*innen abgefragt ⁷⁰. Die Proband*innen, die nicht während der Fahrt abgelenkt wurden, hatten eine realistische Selbsteinschätzung ⁷⁰. Diejenigen, die während der Fahrt durch ein Telefonat abgelenkt wurden, schätzten ihre Fahrsicherheit und Fehleranzahl hingegen wesentlich besser ein, als sie es tatsächlich war ⁷⁰. Daraus schlossen die Autor*innen, dass Ablenkung und Multitasking nicht nur zu einer schlechteren Leistung der primär ausgeführten Aufgabe führen, sondern sich auch negativ auf die korrekte Wahrnehmung der eigenen Leistung auswirken können ⁷⁰.

Zusammenfassend lässt sich sagen, dass der Forschungsstand zu Auswirkungen von Ablenkungen durch Telefonanrufe im Alltag am Beispiel von Autofahren ergibt, dass Ablenkungen zu mehr Fehlern und einer falschen Selbstwahrnehmung führen.

Inwiefern diese Erkenntnisse auf medizinisches Personal und ihre Tätigkeiten übertragen werden können bleibt offen. Die Ergebnisse der Studien zum Autofahren lassen jedoch negative Effekte durch Ablenkungen durch Telefonanrufe auch in Kliniken befürchten. Bisher gibt es nur wenige Studien mit einer entsprechenden Fragestellung. Auf den aktuellen Forschungsstand soll im folgenden Absatz eingegangen werden.

2.4 Forschungsstand zu Auswirkungen von Ablenkungen und Unterbrechungen durch Telefonanrufe im Krankenhaus

Bei der Literaturrecherche wurde deutlich, dass sich bereits viele Studien mit den Effekten von Unterbrechungen und Ablenkungen auf das Arbeiten von medizinischem Personal im Krankenhaus im Allgemeinen beschäftigten ^{71–74}.

In Beobachtungsstudien wurden als Folgen von Ablenkungen und Unterbrechungen von medizinischem Personal eine hohe Arbeitsbelastung ^{5,14–16,23}, Frustration ^{13,23}, ein hohes Stresslevel, ^{9–16} mehr Überarbeitung und Erschöpfung ²⁴ sowie eine schlechtere Behandlungsqualität ⁹ und mehr Fehler in der Medikation ^{19,20} oder während Operationen ^{21,22} festgestellt. Im OP wurden zudem schlechtere Teamarbeit ¹⁴ mit mehr Fehlkommunikation im Team ^{21,75} und längere OP Zeiten ^{76–78} beobachtet.

Nur eine Beobachtungsstudie untersuchte spezifisch den Effekt von Telefonanrufen und stellte einen Zusammenhang zwischen einem fehlerhaft erstellten Befund eines Radiologie Assistenzarztes/ einer Assistenzärztin und Unterbrechungen durch Telefonanrufe während der Befundung fest ⁷⁹. In der Stunde, bevor ein fehlerhafter Befund erstellt wurde, wurden signifikant häufiger Anrufe getätigt und entgegengenommen ⁷⁹.

In den meisten experimentellen Studien zu den Effekten von Ablenkungen und Unterbrechungen von Ärzt*innen wurden die Proband*innen während der Durchführung von simulierten laparoskopischen Eingriffen an einem Laparoskopie Trainer durch Fragen, Mathe- und Konzentrationsaufgaben sowie Hintergrundgeräusche unterbrochen und/oder abgelenkt ^{25,26,28,80–94}. Dabei zeigte sich in einigen Studien, dass die Proband*innen signifikant mehr Fehler machten, wenn sie abgelenkt oder unterbrochen wurden ^{27,28,80,82,87,89,91,92}, in anderen Studien gab es keinen signifikanten Unterschied ^{25,26,81,83,86,90,93,95,96}. In experimentellen Studien zeigte sich zudem, dass die Arbeitserfahrung der Proband*innen einen Einfluss auf die Effekte von Unterbrechungen und Ablenkungen haben kann ^{81,87,97}.

In den experimentellen Studien von Murji et al. und Sujka et al. wurden die Proband*innen während einer simulierten Operation (Salpinxektomie oder Cholezystektomie) durch Pager unterbrochen ^{25,26}. Durch die Pager wurden den Proband*innen zu fiktiven Patientenfällen medizinische Fragen gestellt. In beiden Studien hatten die Ablenkungen zwar keinen Einfluss auf die operativen Leistungen, allerdings beantworteten viele der Proband*innen die Fragen falsch oder mit sicherheitskritischen Konsequenzen für die Patient*innen ^{25,26}.

Yang et al. zeigten, dass sich Telefonanrufe negativ auf die Leistungen der Proband*innen (Medizinstudierende) während der Durchführung einer Laparoskopie auswirken ²⁷. Dies implizierte neben chirurgischen Fehlern auch inadäquate oder inkorrekte Antworten in den Telefongesprächen ²⁷. Die Proband*innen gaben außerdem ein erhöhtes subjektiv empfundenes Stresslevel an ²⁷.

Weigl et al. konnten einen negativen Einfluss von Unterbrechungen durch Telefonanrufe auf die mentale Arbeitsbelastung und technische Leistung ihrer Proband*innen nachweisen²⁸. In der Studie führten Assistenzärzt*innen der Chirurgen unter verschiedenen Bedingungen eine simulierte Vertebroplastie Operation durch. Sie wurden entweder durch einen Telefonanruf abgelenkt oder durch eine akute Verschlechterung des Zustands des operierten Patienten. Bei beiden Ablenkungen erlebten die Proband*innen eine signifikant erhöhte mentale Arbeitsbelastung²⁸. Dabei wurde ein deutlicher Zusammenhang zwischen einer erhöhten mentalen Arbeitsbelastung und technischen Fehlern (Fehlpositionierung des Trocars) festgestellt²⁸.

Zusammenfassend zeigte sich bei der Literaturanalyse, dass in experimentellen Studien Ablenkungen durch Telefonanrufe oder Pager während Operationen zu mehr Fehlern, einer höheren Arbeitsbelastung und Stress führen können.

Insgesamt konnten jedoch nur vier experimentelle Studien, die den Effekt von Telefonanrufen oder Pagern auf die ärztliche Arbeit in Krankenhäusern untersuchten, gefunden werden²⁵⁻²⁸. Diese Studien hatten eine geringe Proband*innenanzahl (maximal 30 Proband*innen)²⁵⁻²⁸ und beschränkten sich auf Ablenkungen während Operationen²⁵⁻²⁸.

Experimentelle Studien zu Ablenkungen durch Telefonanrufe/Pager von Ärzt*innen außerhalb von simulierten Operationen gibt es bislang noch nicht (ausgenommen die vorabveröffentlichten Daten der vorliegenden Arbeit¹).

Dies überrascht, da eine Vielzahl kommunikationsbedingter Unterbrechungen durch Telefone und Pager im direkten Patientenkontakt bei nicht operativen Tätigkeiten auftreten^{4,98} und die Ergebnisse von Beobachtungsstudien nahelegen, dass auch hierbei negative Effekte zu befürchten sind.

Aufgrund der geringen Anzahl von experimentellen Studien mit geringer Proband*innenzahl entschieden wir uns daher eine größere experimentelle Studie durchzuführen, um die Effekte von Unterbrechungen und Ablenkungen durch Telefonanrufe weiterführend zu untersuchen. Im Gegensatz zu den bisherigen experimentellen Studien sollten in der vorliegenden Studie die Effekte von Telefonanrufen auf alltägliche ärztliche Tätigkeiten untersucht werden.

In der Methodik und dem Studiendesign der bisherigen Studien fiel außerdem auf, dass Faktoren, die einen Einfluss auf den Effekt von Unterbrechungen haben, wie beispielsweise Berufserfahrung, keine Berücksichtigung fanden. Auch dies sollte in der vorliegenden Studie ergänzend untersucht werden.

Der Effekt von Telefonanrufen auf die Durchführung alltäglicher ärztlicher Tätigkeiten wurde in der vorliegenden Arbeit im Rahmen von OSCE-Prüfungen (objective structured clinical examination) durchgeführt, darauf soll im folgenden Bezug genommen werden.

2.5 OSCE-Prüfungen in der medizinischen Lehre

OSCE-Prüfungen werden schon seit rund 40 Jahren als Lehr- und Prüfmittel in der medizinischen Lehre eingesetzt. 1975 wurde bereits erstmals der Einsatz einer OSCE-Prüfung an einer medizinischen Fakultät durch Harden et al. beschrieben^{99,100}.

Bei einer OSCE-Prüfung rotieren die Prüflinge durch verschiedene Stationen, in denen sie in einem begrenzten Zeitraum unterschiedliche Aufgaben lösen müssen. Im originalen OSCE nach Harden werden Stationen, in denen praktische Fertigkeiten geprüft werden, mit Stationen abgewechselt, in welchen Multiple Choice Fragen zu der vorherigen Aufgabe beantwortet werden. Die Prüflinge werden außerdem von Prüfer*innen beobachtet und mittels Checklisten bewertet. In den praktischen Stationen setzten Harden et al. u.a. Schauspieler*innen als Simulationspatient*innen ein¹⁰⁰.

Der erfolgreiche Einsatz von standardisierten Patient*innen (geschulte Schauspieler*innen) in der Prüfung von praktischen Fähigkeiten von Medizinstudierenden wurde bereits 1964 von Barrows et al. beschrieben¹⁰¹. Schauspielpatient*innen sind demnach schon seit langem in der Lehre und in Prüfungen von Medizinstudierenden etabliert und werden in bestimmten Situationen echten Patient*innen vorgezogen. So kann gewährleistet werden, dass Studierenden in Prüfungen auf vergleichbare Art und Weise ein Fall präsentiert wird. Außerdem können sie so in einem geschützten Raum lernen¹⁰². Die Reliabilität und Validität von Prüfungen mit standardisierten Patient*innen wird als positiv eingeschätzt und bewertet¹⁰³. Die Schauspielpatient*innen können bei entsprechendem Training eine gute und realistische Simulation ermöglichen¹⁰³.

OSCE-Prüfungen haben sich in den letzten Jahrzehnten stark verbreitet und sind in der medizinischen Lehre als valides und reliables Prüfungs- und Lehrmittel von klinisch praktischen Kompetenzen etabliert. Inzwischen werden OSCE-Prüfungen weltweit sowohl für Medizinstudierende, Ärzt*innen, als auch anderes medizinisches Personal eingesetzt^{104–116}. In Deutschland wurden OSCE-Prüfungen in den vergangenen Jahren an verschiedenen medizinischen Fakultäten durchgeführt¹¹⁷. Im Masterplan 2020 wurde u.a. festgelegt, dass in Zukunft an allen deutschen Universitäten ein einheitliches OSCE als Teil der dritten staatlichen Prüfung am Ende des Medizinstudiums eingeführt werden soll¹¹⁸.

Auch in anderen Studien wurden OSCE-Prüfungen genutzt, um das Verhalten von medizinischem Personal unter außergewöhnlichen Arbeitsbedingungen zu untersuchen. Srisarajivakul et al. untersuchten in zwei unterschiedlichen OSCE Szenarien, wie Assistenzärzt*innen im ersten Jahr ihrer Ausbildung in der Lage sind ihre ärztliche Tätigkeit sorgfältig weiterzuführen, wenn sie mit störendem Verhalten und zwischenmenschlich herausfordernden Situationen konfrontiert wurden¹⁰⁴. Auch Hochberg et al. nutzten OSCE-

Prüfungen um die Professionalität und Kommunikationskompetenz von chirurgischen Assistenzärzt*innen zu evaluieren und Lernerfolge zu prüfen ¹¹⁹.

Wir nutzten daher die OSCE Prüfung als etablierte Prüfungsform, um die Effekte von Telefonanrufen auf alltägliche ärztliche Aufgaben zu testen. Wie die Datenlage zeigt, können in OSCE realistische Simulationen erreicht werden. Der Vorteil lag zudem darin, dass hierbei eine möglichst realistische Situation unter verschiedenen Bedingungen getestet werden konnte, ohne die Studie in direkten Patient*innenkontakt durchführen zu müssen. Durch das standardisierte Vorgehen konnten die Proband*innen die gleiche ärztliche Tätigkeit mit und ohne Ablenkung durchführen und die Leistung verglichen werden. Ein weiterer Vorteil bestand zudem darin, dass durch die Nutzung der bestehenden Strukturen der OSCE-Prüfungen eine große Anzahl von Proband*innen eingeschlossen werden konnte.

2.6 Fragestellungen und Ziel der Arbeit

Ziel der Arbeit war es, Auswirkungen von Telefonanrufen auf die Leistung von medizinischen Personal bei alltäglichen ärztlichen Tätigkeiten in einem experimentellen Studiendesign zu untersuchen und hiermit eine Forschungslücke zur Patient*innenensicherheit zu füllen.

Aufgrund der in den vorherigen Kapiteln skizzierten Problematik ergaben sich folgende Fragestellungen der Arbeit:

I Haben Telefonanrufe einen negativen Effekt auf die erbrachte Leistung der Proband*innen in einer OSCE Prüfung?

II Haben Eigenschaften der Proband*innen einen Einfluss auf den Effekt der Unterbrechungen durch Telefonanrufe? Können Erfahrungen im Umgang mit Unterbrechungen und Ablenkungen das Risiko für negative Effekte durch Ablenkungen durch Telefonanrufe reduzieren?

III Wie wirken sich Telefonanrufe auf die Bearbeitungsdauer der primären Tätigkeit aus?

IV Wie ist die qualitative Durchführung der Telefongespräche, die Verarbeitung von darin genannten Informationen und die Gedächtnisleistung der Proband*innen?

V Können Handlungsempfehlungen für die Praxis und die Lehre abgeleitet werden, um die Patient*innensicherheit zu verbessern?

3. Material und Methoden

3.1 Studiendesign

3.1.1. Kollektiv

Die Studie wurde im Sommersemester 2019 im Rahmen der OSCE-Prüfungen I und II an der Medizinischen Fakultät der Universität zu Köln durchgeführt.

Die OSCE-Prüfungen sind fester Bestandteil des Curriculums im ersten (OSCE I) und fünften (OSCE II) klinischen Semester.

Im Rahmen der vorliegenden Arbeit wurden in OSCE I und II in einer regulär stattfindenden Station sowie in einer zusätzlichen Studienstation Daten erhoben.

Die Prüfungsteilnehmer*innen wurden in einer Einführungsveranstaltung über die Studie und eine freiwillige Teilnahme informiert. Eine schriftliche Einverständniserklärung wurde jedem Probanden/ jeder Probandin ausgehändigt und unterschrieben eingesammelt. Die Proband*innen füllten zusätzlich einen Fragebogen (siehe Anhang) aus. Zusätzlich wurde an den Tagen der Prüfungen, im Rahmen der Anmeldung und Einweisung vor Ort erneut über die Studie und eine mögliche Teilnahme informiert. Außerdem wurde dort überprüft, ob von jedem eingeschlossenen Studierenden eine ausgefüllte Einverständniserklärung und ein ausgefüllter Fragebogen vorlagen. Um alle Studierende eines Semesters zu prüfen, dauerten die OSCE-Prüfungen drei Tage. Die Studierenden wurden zuvor in Gruppen eingeteilt und erfuhren im Vorhinein ihren Prüfungstag.

Wenn ein Studierender nicht an der Studie teilnehmen wollte, hat er während der Studienstation pausiert und im Raum der Studienstation auf den Beginn der nächsten Station gewartet.

Für die Arbeit liegt ein positives Votum der Ethikkommission der Medizinischen Fakultät unter der Antragsnummer 19-1327 vor.

3.1.2. Time series, within-group Design

Bei der Studie handelt es sich um eine experimentelle Studie mit einem time series, within-group Design ¹²⁰.

Dies bedeutet, dass der Effekt von Unterbrechungen durch Telefonanrufe innerhalb der gleichen Gruppe untersucht wurde. Es gab keine Kontrollgruppe, stattdessen fungierte die Gruppe als Kontrolle für sich selbst ¹²⁰.

Die Proband*innen haben im Rahmen der OSCE-Prüfung in zwei Stationen die gleiche Aufgabe bearbeitet. In OSCE I handelte es sich um das Legen eines Venenverweilkatheters und bei OSCE II um ein Anamnesegespräch, eine körperliche Untersuchung eines Schauspielpatienten/ einer Schauspielpatientin, das Stellen einer Verdachtsdiagnose und das Aussprechen einer Therapieempfehlung.

Alle Proband*innen haben die jeweilige Aufgabe einmal ohne Unterbrechung regulär durchgeführt und einmal mit Unterbrechung durch zwei Telefonanrufe (im Folgenden Studienstation genannt). Die Reihenfolge wurde zufällig festgelegt. Da jeder Proband/ jede Probandin nacheinander unter beiden Bedingungen die Prüfungsaufgabe bearbeitete, handelte es sich um ein time series Design ¹²⁰.

Folgende abhängigen Variablen wurde erhoben:

- Leistung der Proband*innen, bewertet mit einer Checkliste
- Korrektes Wiedergeben von Informationen aus den Telefongesprächen ja/nein
- Prüfungsaufgabe im vorgegebenen Zeitraum abgeschlossen ja/nein

Folgende unabhängige Variablen wurden erhoben:

- Experimentelle Variable: Ablenkung durch Telefonanrufe ja/nein
- Subjektvariablen: demographische Daten, Muttersprache, Nationalität, Variablen, die mit Vorerfahrungen der Proband*innen mit Unterbrechungen assoziiert sein könnten. (Siehe auch Fragebogen im Anhang)

3.1.3. Experimentelles Vorgehen

Für die OSCE-Prüfungen wurden die Räumlichkeiten des Kölner Interprofessionellen Skills Lab und Simulationszentrums genutzt. Die Positionierung der Studienstation kann den beiden folgenden Abbildungen entnommen werden. Die OSCE I Prüfung umfasste acht Stationen, die OSCE II Prüfung 14 Stationen. Die Stationen befanden sich in nummerierten, benachbarten Räumen. Die Zuteilung zur ersten Station erfolgte zufällig nach Prüfungs-ID (Buchstabenkombination).

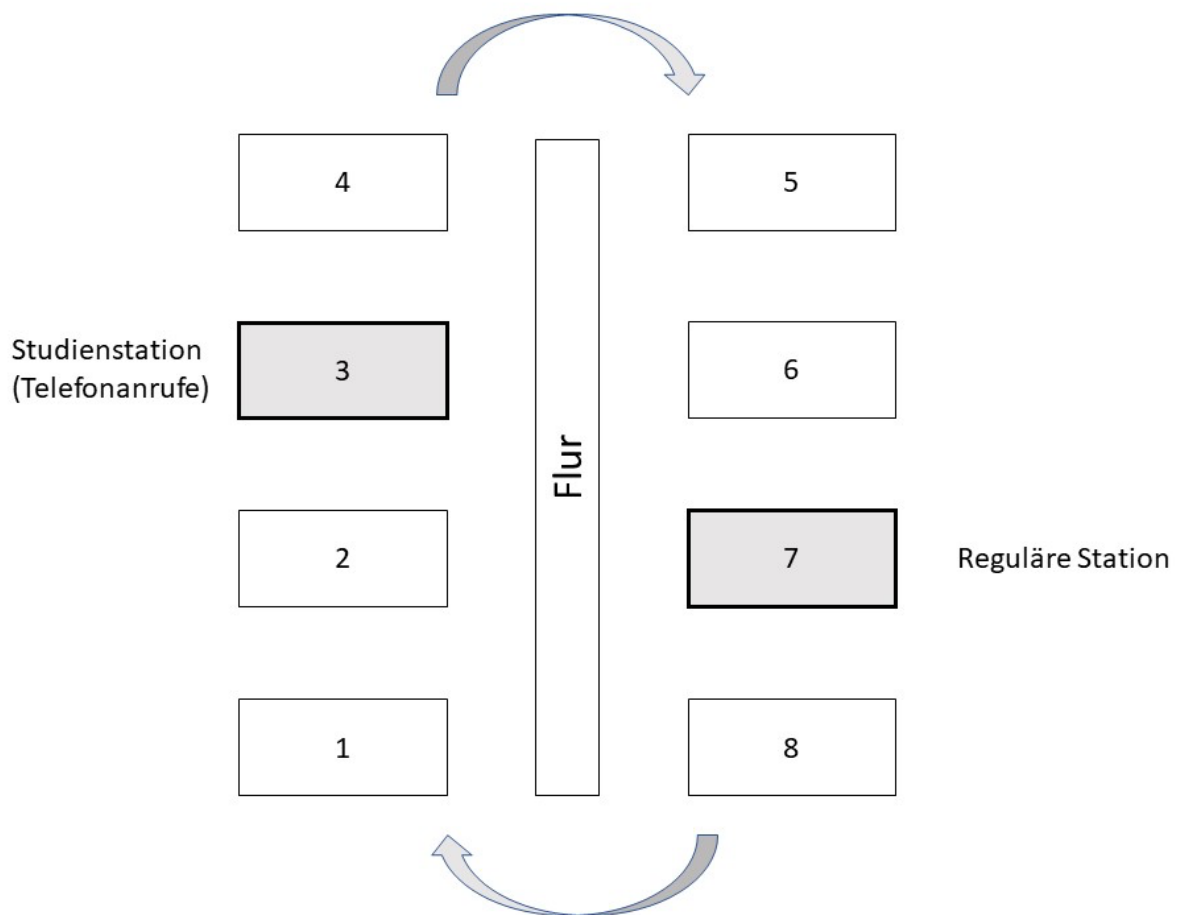


Abbildung 1, Aufbau OSCE I

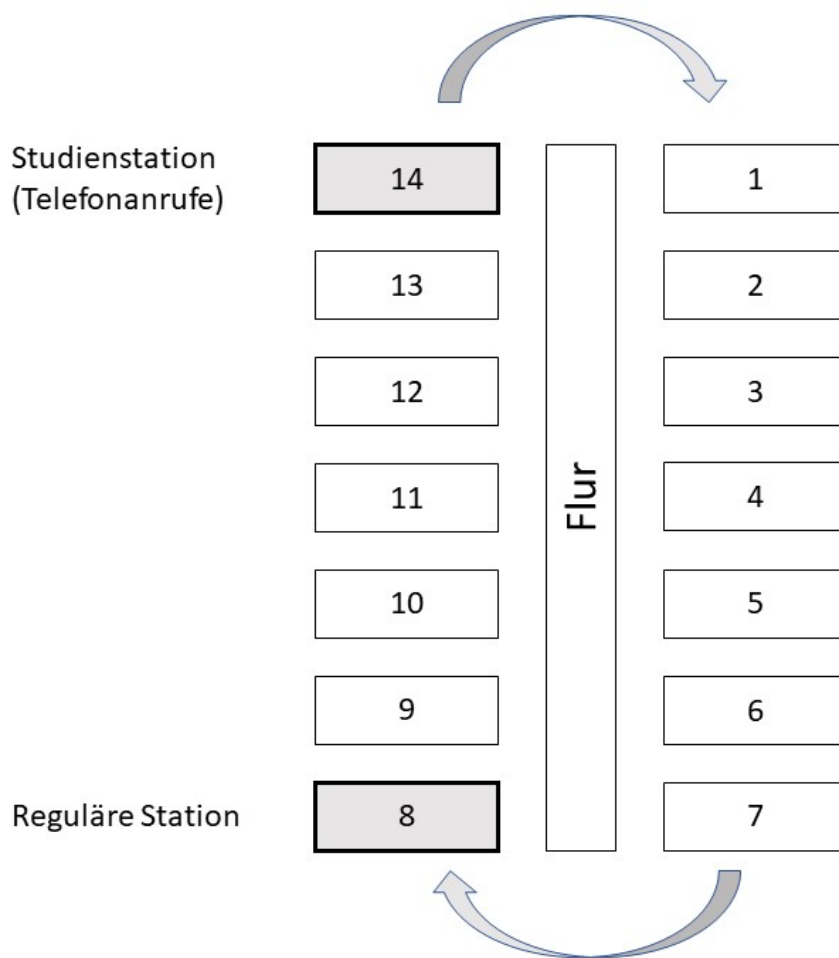


Abbildung 2, Aufbau OSCE II

Durch die Positionierung haben bei OSCE I 50% der Proband*innen zuerst die reguläre Station durchlaufen und 50% zuerst die Studienstation.

Bei OSCE II haben 43,5% der Proband*innen zuerst die Studienstation durchlaufen und 56,5% zuerst die reguläre Station durchlaufen. Aufgrund der örtlichen Gegebenheiten und den logistischen Anforderungen der anderen Stationen konnte keine Aufteilung in jeweils 50% erfolgen.

3.1.4. Aufgabenstellung

Die Aufgabenstellungen der Studienstationen werden regelhaft in den OSCE-Prüfungen angewendet. Diese können den Fallvignetten im Anhang entnommen werden.

In OSCE I sollten die Proband*innen einen Venenverweilkatheters legen und bei OSCE II ein Anamnesegespräch und eine körperliche Untersuchung eines Schauspielpatienten/ einer Schauspielpatientin durchführen sowie eine Verdachtsdiagnose stellen und eine Therapieempfehlung geben.

In OSCE II wurden drei Erkrankungen geprüft: Cholezystitis, Appendizitis, Divertikulitis. Die Erkrankungen wurden nach jeder Gruppe rotiert. Auch innerhalb einer Gruppe wurden in der Studienstation und in der regulären Station unterschiedliche Erkrankungen geprüft. (S.a. Tabelle im Anhang mit Einteilung der Szenarien, unter 7.3.3.4).

Die Schauspielpatient*innen in der regulären OSCE Station waren von der Medizinischen Fakultät bezahlte, professionelle Schauspieler*innen. Die Schauspieler*innen in der Studienstation waren Medizinstudierende (Doktorand*innen der Arbeitsgruppe). Um eine Vergleichbarkeit gewährleisten zu können, wurden alle Schauspieler*innen in einer vorherigen gemeinsamen Schulung angeleitet und erhielten ein Skript, welches Informationen zu der Krankengeschichte, zu den typischen Untersuchungsbefunden und Symptomen enthielt.

Auf der Fallvignette der Studienstation war zusätzlich vermerkt, dass die Proband*innen in ihrer Rolle als diensthabender Arzt/ diensthabende Ärztin Telefonanrufe über ein Diensttelefon beantworten mussten.

3.1.5. Zeitlicher Ablauf

Den Abbildungen 3 und 4 kann der zeitliche Ablauf der Studienstationen mit den Telefonanrufen entnommen werden.

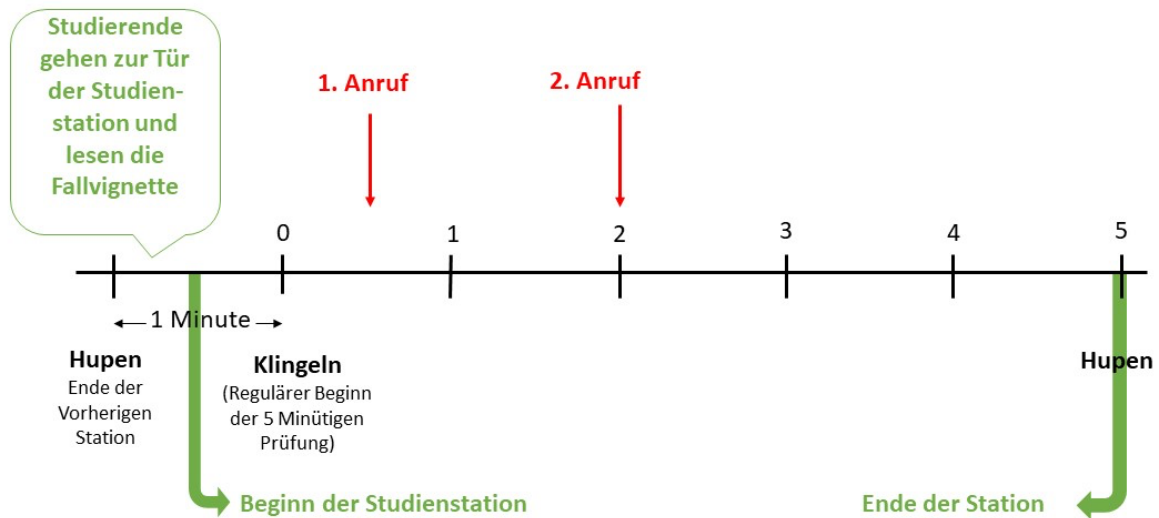


Abbildung 3, Zeitlicher Ablauf Studienstation OSCE I

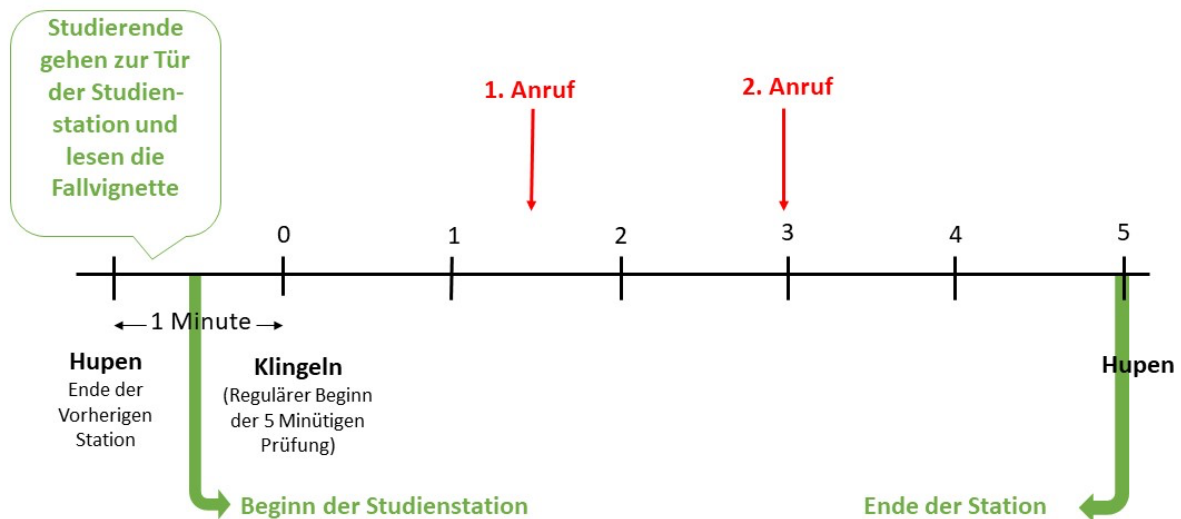


Abbildung 4, Zeitlicher Ablauf Studienstation OSCE II

Um Prüfungsbeginn und -ende zu signalisieren, wurden über Lautsprecher Töne abgespielt. Zu Beginn der Prüfung ertönte ein Hupen, jeder Prüfling stand vor einer Station und las die an der Tür befestigte Fallvignette mit der Aufgabenstellung. In den regulären Stationen ertönte nach einer Minute ein Klingeln und die Prüflinge betraten die Station.

In den Studienstationen begangen die Proband*innen früher mit der Prüfung. Durch den früheren Beginn erfolgte ein Zeitausgleich für die ca. 30 Sekunden andauernden Unterbrechung durch die Telefonanrufe. Über den früheren Beginn wurden die Proband*innen im Einführungsgespräch, während der Anmeldung am Prüfungstag und durch die Fallvignette informiert.

Den Proband*innen wurde beim Betreten der Studienstation ein Telefon ausgehändigt. Die Anrufe wurden von Doktorand*innen der Arbeitsgruppe durchgeführt. Zu diesem Zweck wurden diese zuvor geschult und erhielten ein Skript. Sie hielten sich in einem separaten Raum auf und tätigten zu vorher festgelegten Zeitpunkten die Anrufe.

Die Zeitpunkte für die Anrufe wurden in vorherigen Probedurchläufen mit dem Ziel festgelegt, dass die Proband*innen sich in einer bestimmten Phase der Aufgabenbearbeitung befanden. Bei OSCE I erfolgte der erste Anruf 30 Sekunden nach dem Klingeln, um die Proband*innen beim Richten der Materialien für die Venenpunktion zu unterbrechen. Der zweite Anruf erfolgte zwei Minuten nach dem Klingeln, um die Proband*innen kurz vor, während oder nach der Punktion zu unterbrechen.

Bei OSCE II erfolgte der erste Anruf 1,5 Minuten nach dem Klingeln, um die Proband*innen bei der Anamnese zu unterbrechen. Der zweite Anruf erfolgte drei Minuten nach dem Klingeln, um die Proband*innen bei der körperlichen Untersuchung zu unterbrechen.

Nach fünf Minuten signalisierte ein Hupen das Ende der Bearbeitungszeit. Die Prüflinge mussten nun innerhalb einer Minute die Station verlassen und die Fallvignette der nächsten Station lesen. Dies wiederholte sich, bis alle Studierenden einer Prüfungsgruppe alle Stationen durchlaufen haben. Eine OSCE I Prüfung dauert ca. 50 Minuten, eine OSCE II Prüfung ca. 85 Minuten.

3.1.6. Inhalt der Telefonanrufe

Der Inhalt der Telefonanrufe wurde in Zusammenarbeit mit den Ärzt*innen aus der Klinik und Poliklinik der Allgemein-, Viszeral-, Thorax- und Transplantationschirurgie der Uniklinik Köln aus der beteiligten Arbeitsgruppe konzipiert, um eine möglichst realistische Ablenkung zu erreichen. In dem ersten Anruf erhielten die Proband*innen spezifische Informationen zu einem fiktiven Patienten/ einer fiktiven Patientin, in dem zweiten Anruf mussten Details aus dem ersten Anruf wiedergegeben werden.

In OSCE I lautete der Anruf; „Hallo, Pfleger*in X hier. Der bekannte Patient Herr/Frau X für die OP morgen hat einen Blutzuckerwert von X. Ihm geht es gut. Ich wollte nur Bescheid sagen, dass ich ihm so wie immer nach dem Standardschema Insulin spritze.“

In OSCE II erfolgte eine andere Variante des Anrufs: „Hallo, Pfleger*n X hier. Der bekannte Patient Herr/Frau X für die OP morgen hat einen Blutzuckerwert von X. Ihm geht es gut. Soll ich so wie immer nach dem Standardschema Insulin spritzen?“.

Es wurde festgelegt, dass bei Zögern oder Fragen des Probanden/ der Probandin der Anruf nach spätestens 30 sec mit der Aussage, der Anrufende habe keine Zeit mehr, beendet wurde. Im Vergleich zu OSCE I sollten die Proband*innen bei OSCE II auch eine kurze Einschätzung abgeben, ob eine Insulingabe sinnvoll ist. Dadurch sollte der Ablenkungseffekt noch verstärkt werden. In OSCE I wurde von dem Einfordern der aktiven Zustimmung der Proband*innen zur Insulingabe abgesehen, da nicht gewährleistet werden konnte, dass diese im 5. Semester über ausreichendes Wissen über die Insulintherapie bei Diabetes Mellitus verfügten, um eine medizinisch fundierte Entscheidung treffen zu können.

Der zweite Anruf war bei OSCE I und OSCE II identisch: „Hallo Oberärzt*in X hier, die Anästhesie will für den/die Patient/in der/die morgen die OP hat den Blutzucker wissen. Ach, wie hieß er/sie denn nochmal? Und wie war der Blutzucker?“ Die Antworten der Proband*innen wurden notiert. Wenn sie keine Antwort wussten, wurde ihnen gesagt, dass der Anrufende jemand anderen fragen würde.

Zur Reduktion einer Treatmentdiffusion auf Grund von Informationsaustausch unter den Proband*innen wurden die Namen der Patient*innen und die Blutzuckerwerte in den verschiedenen Prüfungsgruppen variiert. Es wurde zwischen den Namen Herr Weber, Herr Mutlu, Frau Fischer und den Blutzuckerwerten 205, 210, und 220 mg/dL variiert (s.a. Anhang).

3.2 Dokumentation und Bewertung

Die Leistungen in der regulären Station wurden durch einen Prüfer/ eine Prüferin der Medizinischen Fakultät mit einem offiziellen Prüfungsbogen (Checkliste) erfasst.

Für die Studie wurden die Checklisten der Universität überarbeitet und zur Bewertung der Leistungen der Proband*innen in beiden Stationen eingesetzt. Die dafür eingesetzten Prüfer*innen waren Doktorand*innen der Arbeitsgemeinschaft, welche zuvor in der Bewertung und Nutzung des Prüfungsbogens geschult wurden.

Die Checklisten finden sich in den folgenden Tabellen (Tabelle 1 und 2). Es konnten insgesamt maximal 10 Punkte erreicht werden. Auf der Checkliste der Studienstation wurde zusätzlich vermerkt, ob die Proband*innen den Blutzuckerwert und Patientennamen korrekt wiedergeben konnten und ob diese die Station vor Ablauf der Zeit abschließen konnten.

Die Leistungen in der normalen Prüfungsstation sind in die offizielle Bewertung der OSCE-Prüfung eingeflossen. Die Leistungen in der Studienstation wurden nur für die Studie beurteilt und sind nicht in die offizielle Bewertung eingeflossen.

In der regulären Station wurden die Proband*innen demnach durch einen offiziellen Prüfer/ eine offizielle Prüferin der Universität und zusätzlich durch einen Prüfer/ eine Prüferin der Studie beurteilt.

Bei OSCE I befanden sich die Prüfer*innen im gleichen Raum wie die Proband*innen. In OSCE II befanden sich die Prüfer*innen in der regulären Station im Nebenraum und konnten über eine einseitig verspiegelte Glasscheibe und Kopfhörer die Prüfung verfolgen. In der Studienstation befand sich der Prüfer/ die Prüferin aus logistischen Gründen im gleichen Raum.

Die Studierende hatten Prüfungs-IDs, anhand derer die Daten anonymisiert aufgenommen werden konnten. Die Prüflinge trugen während der Prüfung an ihren Kitteln Aufkleber mit ihrer Prüfungs-ID.

Die Prüfungsbögen wurden über Tablets ausgefüllt. Die Datenverarbeitung erfolgte über ein von der Universität entwickeltes Programm. Die Checklisten wurden zuvor in die Datenbank des Programms eingepflegt und auf einem Server des Regionalen Rechenzentrum der Universität zu Köln gespeichert. Während der Prüfung wurden die Checklisten über das Uniklinik Netzwerk ukknow oder das Universitätsnetz Eduroam (beide nur über einen Account nutzbar) auf den Tablets abgerufen, mit einer Prüfungs-ID gekennzeichnet und ausgefüllt. Nach der Eingabe wurden die Checklisten mit einem Zeitstempel versehen und an die Datenbank übermittelt. Die ausgewerteten Checklisten konnten anschließend heruntergeladen werden.

Für den Fall, dass technische Probleme ein digitales Ausfüllen der Prüfungsbögen unmöglich gemacht hätten, standen in jedem Prüfungsraum ausgedruckte Exemplare der Prüfungsbögen zur Verfügung. Ein Einsatz dieser Bögen war nicht nötig.

Tabelle 1, Checkliste OSCE I

Arbeitsschritt	Richtig	Teilweise	Falsch
Richten des Materials			
Venenverweilkatheter	0.16		0
Stauschlauch	0.16		0
Desinfektionsmittel	0.16		0
Pflaster	0.16		0
Tupfer	0.16		0
Abwurf	0.20		0
Handschuhe anziehen (Teilweise wenn ohne Händedesinfektion)	1.0	0.5	0
Anlegen des Stauschlauchs (Teilweise wenn zu lange vor der Punktion)	0.5	0.25	0
Ablegen des Stauschlauchs (Teilweise wenn nach Herausziehen der Nadel)	0.5	0.25	0
Desinfektion der Punktionsstelle (kein Nachtasten, nicht mit Tupfer abwischen)	0.5		0
30 Sekunden Einwirkzeit der Desinfektion	0.5		0
Zügiges Punktieren (neue Viggo bzw. neuer Versuch: max. "teilweise")	0.5	0.25	0
Sauberes Punktieren der Vene (nein wenn Nadel unsteril / gebogen)	0.5		0
Abwurf der Punktionsnadel (teilweise wenn Nadel aufs Tablett und später in den Abwurf; "nein" wenn mit offener Nadel durch den Raum gelaufen)	1.0	0.5	0
Fixieren der Viggo	1.0	0,5	0
Anschließen und Kontrollieren der Infusion	1.0	0,5	0
Patienten über die Punktion informieren	0.5		0
Strukturiertes Arbeiten	1.0	0,5	0
Händedesinfektion nach Ausziehen der Handschuhe	0.5		0

Tabelle 2, Checkliste OSCE II

Arbeitsschritt	Richtig	Teilweise	Falsch
<i>Anamnese</i>			
Vollständige Exploration der Schmerzen: was, wie, seit wann, wo, wie oft	1	0,5	0
Vollständige Exploration der vegetativen Symptome: Übelkeit? Erbrechen? Stuhlgang? Fieber?	1	0,5	0
Sonstige Anamnese: Vorerkrankungen? Vor-Operationen?	1	0,5	0
Medikamentenanamnese? Allergien?	1	0,5	0
<i>Körperliche Untersuchung:</i>			
Lagerung liegend, Arme seitlich des Pat.	0,8		0
Kommentieren der Inspektion	0,8		0
Auskultation in allen 4 Quadranten	0,8		0
Perkussion	0,8		0
Palpation:			
Beginn in schmerzloser Region, oberflächliche und tiefe Palpation	0,4		0
Je nach Szenario: Sigmadivertikulitis: Überprüfen des Loslassschmerz Cholezystitis: Murphy Zeichen Appendizitis: Überprüfen des Loslassschmerz (Blumberg-Zeichen),/Mc Burney/ Lanz-Punkt/ Psoas-Zeichen	0,4		0
<i>Verdachtsdiagnose:</i>			
Benennen einer Verdachtsdiagnose Erklären des weiteren Vorgehens (weitere Diagnostik, Therapie)	1	0,5	0
<i>Umgang mit Patient</i>			
Professionelle Kontaktaufnahme Kommuniziert in Patientengerechter Sprache Zeigt Empathie Hygienisches Arbeiten	1	0,5	0

3.3 Materialien

3.3.1. OSCE I

Die in den OSCE I Prüfungen eingesetzten Materialien sind der Tabelle 3 zu entnehmen.

Tabelle 3, Materialien OSCE I

Eingesetztes Material OSCE I
- Patientenliege oder -bett - Drehhocker - Resusci Anne Torso von Laerdal - Punktions-Arm (Standard Venipuncture Arm – Light von Limbs & Thinks) mit Kunstblut - Kissen und Bettdecke - Nachttisch - Infusionsständer, 0,9% NaCl Infusion, Infusionsbesteck - Handschuhe Größen S, M, L - Händedesinfektionsmittel - Abwurfbehälter - Tablett - Schubladenschrank mit verschiedenen Fächern - Venenverweilkanülen in verschiedenen Größen - Venenverweilkanülen Pflaster - Fixierungspflaster (Rollen) - 5x5 cm Pflaster - Stauschlauch - Hautdesinfektionsmittel - Verschlusskappen - Tupfer - Kompressen 10x10 - Spritzen 2ml - Nadeln - Smartphone als Diensttelefon - Stoppuhr und Smartphone für Anrufer/Anruferin - Tablet und Ladekabel für Prüfer/Prüferin

3.3.2. OSCE II

Die in den OSCE II Prüfungen eingesetzten Materialien sind der Tabelle 4 zu entnehmen.

Tabelle 4, Materialien OSCE II

Eingesetztes Material
- Desinfektionsspender - Handschuhe Größe S, M, L - Untersuchungsliege - Lagerungsrolle - Drehhocker - Stethoskop - Smartphone als Diensttelefon - Stoppuhr und Smartphone für Anrufer/Anruferin - Tablet und Ladekabel für Prüfer/Prüferin

3.4 Aufgaben der beteiligten Doktorand*innen

Die Studie wurde mit Hilfe von Doktorand*innen der Arbeitsgemeinschaft Lehre umgesetzt. Caroline Kroll, Martina Graupner, Moritz Zeeh, Nadja Jakovljevic, Thomas Dratsch und Louisa Roentgen unterstützen die Studie als Prüfer*innen zur Erhebung der Studiendaten sowie als Schauspieler*innen und Anrufer*innen während der Durchführung der Studie. Louisa Roentgen (Autorin der Dissertation) plante und koordinierte zudem die Studienvorbereitung und Durchführung.

3.5 Statistische Auswertung

Zur Berechnung der benötigten Fallzahl wurde eine statistische Power-Analyse durchgeführt. Mit einem Alpha von .05 und einer Power von .80 ergab sich eine Fallzahl von mindestens $N = 90$ Probanden für das Auffinden eines mittleren Effektes für Intragruppenvergleiche. In Bezug auf Korrelationen wurde eine Fallzahl von $N = 34$ für das Auffinden eines mittleren Effektes ermittelt.

Alle Auswertungen erfolgten mit der Statistiksoftware Statistical Package for the Social Sciences (SPSS, Version 25; IBM, 2017).

Für die Berechnung des primären Outcome wurden die durchschnittlich erreichten Gesamtpunktzahlen der Proband*innen in der Studienstation mit der regulären Station mittels t Test verglichen. Außerdem erfolgte ein Vergleich nur unter den Proband*innen, die rechtzeitig in der Studienstation und regulären Station fertig wurden.

Es erfolgte zudem eine Aufschlüsselung nach einzelnen Items der Checkliste und ein Vergleich der erreichten Punktzahl mittels t Test.

Außerdem erfolgte eine Auswertung, ob die Informationen aus dem Anruf korrekt wiedergegeben werden konnten.

Zudem eine Analyse der Variablen: Ausbildung im medizinischen Bereich, Muttersprache, Sport, Geschwister und Videospiele, ebenfalls mit t Test.

4. Ergebnisse

4.1 Alter und Geschlecht der Proband*innen

4.1.1. OSCE I

161 Medizinstudierende konnten im Rahmen ihrer OSCE I Prüfung in die Studie eingeschlossen werden. Davon waren 76 Männer und 85 Frauen. Das mittlere Alter betrug 23,7 Jahre. Der/die jüngste Proband*in war 20 Jahre alt und der/die älteste Proband*in war 33 Jahre alt.

4.1.2. OSCE II

Im Rahmen ihrer OSCE II Prüfung nahmen 147 Medizinstudierende an der Studie teil. Davon waren 61 Männer und 86 Frauen. Das mittlere Alter lag bei 25,7 Jahren. Der/ die jüngste Proband*in war 22 Jahre alt, der/die älteste Proband*in 43 Jahre.

4.2 Effekt der Telefonanrufe

Um zu prüfen, ob die Leistung der Proband*innen in der OSCE-Prüfung schlechter war, wenn sie dabei durch die Telefonanrufe unterbrochen wurden, wurde ihre erreichte Punktzahl mit ihrer Punktzahl in der Station ohne Telefonanrufe mittels gepaarten t-Test verglichen.

4.2.1. Vergleich der erreichten Punktzahl, OSCE I

In OSCE I hatten die Proband*innen durchschnittlich eine signifikant höhere Anzahl von Punkten in der normalen OSCE Station ohne Telefonanrufe ($M = 6.44$, $SD = 1.77$) als in der OSCE Station mit Telefonanrufen ($M = 5.95$, $SD = 1.69$), $t(160) = 2.96$, $p = .004$, $d = 0.23$. Daraus ergibt sich, dass die Proband*innen bei der Prüfung ein besseres Ergebnis erzielten, wenn sie nicht durch die Telefonanrufe abgelenkt wurden.

4.2.2. Vergleich der erreichten Punktzahl, OSCE II

In OSCE II war ebenfalls ein signifikanter Unterschied in der durchschnittlichen Gesamtpunktzahl zwischen der normalen OSCE Station ohne Telefonanrufe ($M = 7.59$, $SD = 1.20$) und der OSCE Station mit Telefonanrufen ($M = 6.84$, $SD = 1.23$), $t(146) = 6.55$, $p < .001$, $d = .55$. Demnach absolvierten die Proband*innen die Prüfung mit einem besseren Ergebnis, wenn sie nicht durch Telefonanrufe abgelenkt wurden.

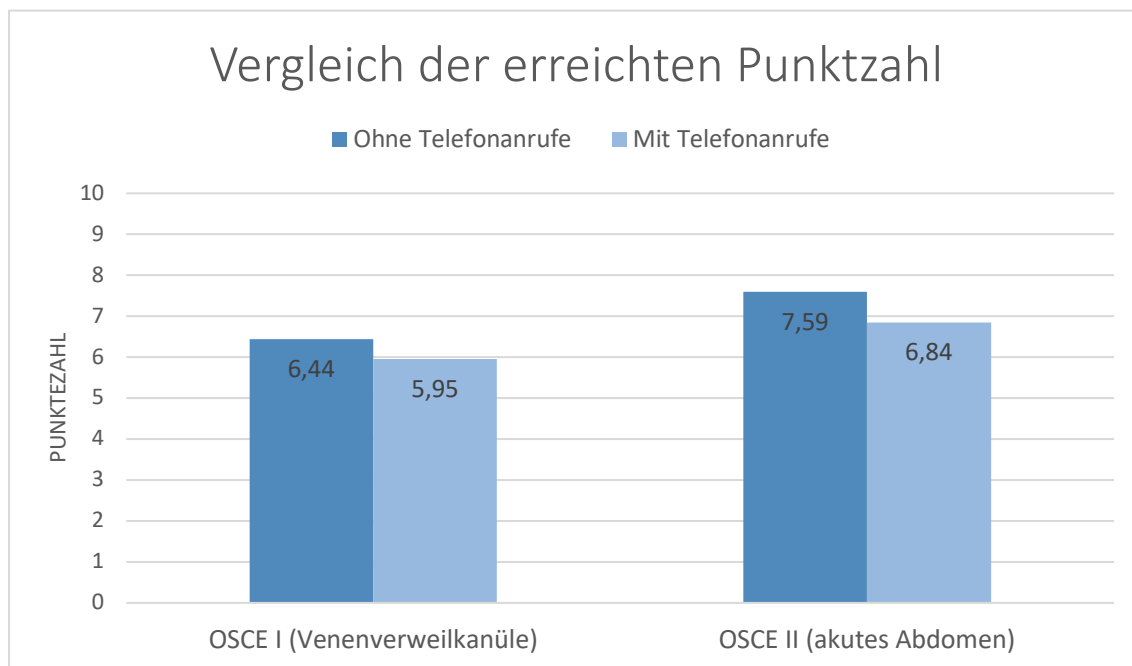


Abbildung 5, Vergleich der erreichten Punktzahl OSCE I und OSCE II ¹

In Abbildung 5 wird der Unterschied in den erreichten Punktzahlen in OSCE I und OSCE II deutlich. Die Abbildung orientiert sich an der Abbildung im vorab veröffentlichten Paper ¹.

4.2.3. Bearbeitung der Aufgabe im vorgegebenen Zeitrahmen

Um auszugleichen, dass die Telefonanrufe die Zeit zur Bearbeitung der OSCE Aufgabe minimierten, wurde den Proband*innen in den Studienstationen mit Telefonanrufen mehr Zeit zur Verfügung gestellt, indem sie früher beginnen durften.

Um dennoch sicher zu sein, dass ein Unterschied in der erreichten Punktzahl in den beiden Stationen nicht darauf zurückzuführen ist, dass die Proband*innen aus Zeitmangel in der Station mit Telefonanrufen nicht fertig wurden, wurde die Anzahl an Proband*innen, die die jeweiligen Stationen nicht in der vorgegebenen Zeit abschlossen, miteinander verglichen.

In OSCE I wurden 75.8% der Proband*innen (122 von 161) in der normalen OSCE Station ohne Telefonanrufe nicht fertig, in der Station mit Telefonanrufen wurden 75.8% der Proband*innen (122 von 161) nicht fertig.

In OSCE II wurden 14.3% der Proband*innen (21 von 147) in der normalen Station ohne Telefonanrufe nicht fertig und 23.1% der Studierenden (34 von 147) in der Station mit Telefonanrufen.

Wenn nur die Proband*innen, die beide Stationen in der vorgegebenen Zeit beendeten, untereinander verglichen werden, ist dennoch ein signifikanter Unterschied in der erreichten Punktzahl festzustellen.

Im Vergleich hatten die Proband*innen in OSCE I, die beide Stationen in der vorgegeben Zeit beendeten ($n = 13$), in der normalen OSCE Station ohne Telefonanrufe ($M = 8.48$, $SD = 0.89$) signifikant mehr Punkte als in der OSCE Station mit Telefonanrufen ($M = 7.39$, $SD = 0.91$), $t(12) = 3.84$, $p = .002$, $d = 1.07$.

In OSCE II hatten die Proband*innen, die in beiden Stationen rechtzeitig fertig wurden ($n = 101$), signifikant mehr Punkte in der normalen OSCE Station ohne Telefonanrufe ($M = 7.61$, $SD = 1.16$) als in der OSCE Station mit Telefonanrufen ($M = 6.92$, $SD = 1.23$), $t(100) = 5.21$, $p < .001$, $d = 0.53$.

Daher ist davon auszugehen, dass die Unterschiede in der erreichten Punktzahl zwischen der normalen OSCE Station ohne Telefonanrufe und der OSCE Station mit Telefonanrufen, nicht daran liegen, dass die Proband*innen wegen den Telefonanrufen nicht rechtzeitig fertig wurden.

4.2.4. Unterschiede in der erreichten Punktezahl aufgeschlüsselt nach Items in der Checkliste

4.2.4.1 OSCE I

Tabelle 5 zeigt, wie sich die Bewertung der einzelnen Items der Checkliste in der Studienstation mit Telefonanrufen und in der regulären Station in OSCE I unterschieden. Es zeigt sich, dass die Proband*innen in der Studienstation mit Telefonanrufen signifikant weniger Punkte für das Richten der Materialien (Desinfektionsmittel, Tupfer, Abwurf, Stauschlauch), Handschuhe anziehen, Anlegen des Stauschlauchs und zügiges Punktieren erhielten. Insgesamt wurden in allen Items, außer „Anschließen und Kontrollieren der Infusion“, weniger Punkte in der Studienstation gegeben, jedoch ist der Unterschied nicht bei allen Items signifikant. Die Tabelle orientiert sich an der Tabelle im vorab veröffentlichten Paper ¹.

Tabelle 5, Unterschiede in der Bewertung der Checklisten Items zwischen der regulären Station und der Studienstation in OSCE I ¹

Arbeitsschritt	Mittelwertdifferenz	p
Richten des Materials		
Venenverweilkatheter	-0.00199	.158
Stauschlauch	-0.01292	.001
Desinfektionsmittel	-0.01391	<.001
Pflaster	0.00099	.797
Tupfer	-0.01491	.002
Abwurf	-0.0248	.002
Handschuhe anziehen (Teilweise wenn ohne Händedesinfektion)	-0.0776	.01
Anlegen des Stauschlauchs (Teilweise wenn zu lange vor der Punktion)	-0.0295	.016
Ablegen des Stauschlauchs (Teilweise wenn nach Herausziehen der Nadel)	-0.02019	.366
Desinfektion der Punktionsstelle (kein Nachtasten, nicht mit Tupfer abwischen)	-0.0124	.451
30 Sekunden Einwirkzeit der Desinfektion	-0.0186	.181
Zügiges Punktieren (neue Viggo bzw. neuer Versuch: max. "teilweise")	-0.06832	.001
Sauberes Punktieren der Vene (nein wenn Nadel unsteril / gebogen)	-0.0155	.494

Abwurf der Punktionsnadel (teilweise wenn Nadel aufs Tablett und später in den Abwurf; "nein" wenn mit offener Nadel durch den Raum gelaufen)	-0.0683	.144
Fixieren der Viggo	-0.0373	.436
Anschließen und Kontrollieren der Infusion	0.0031	.942
Patienten über die Punktion informieren	-0.0311	.114
Strukturiertes Arbeiten	-0.0342	.26
Händedesinfektion nach Ausziehen der Handschuhe	-0.0124	.451

4.2.4.2 OSCE II

Tabelle 6 zeigt, wie sich die Bewertung der einzelnen Items der Checkliste in der Studienstation mit Telefonanrufen und in der regulären Station in OSCE II unterschieden. Es zeigt sich, dass die Proband*innen in der Studienstation mit Telefonanrufen signifikant weniger Punkte für einen Teil der Anamnese (Vorerkrankungen und -operationen sowie Allergien und Medikamente), Inspektion, Auskultation und Perkussion erhielten.

Nur für die korrekte Lagerung des Patienten/ der Patientin erhielten die Proband*innen in der Studienstation signifikant mehr Punkte als in der regulären Station.

Die Tabelle orientiert sich an der Tabelle im vorab veröffentlichten Paper ¹.

Tabelle 6, Unterschiede in der Bewertung der Checklisten Items zwischen der regulären Station und der Studienstation in OSCE II ¹

Arbeitsschritt	Mittelwertdifferenz	p
<i>Anamnese</i>		
Vollständige Exploration der Schmerzen: was, wie, seit wann, wo, wie oft	0.0136	.573
Vollständige Exploration der vegetativen Symptome: Übelkeit? Erbrechen? Stuhlgang? Fieber?	-0.0408	.287
Sonstige Anamnese: Vorerkrankungen? Vor-Operationen?	-0.1463	<.001
Medikamentenanamnese? Allergien?	-0.1531	<.001
<i>Körperliche Untersuchung:</i>		
Lagerung liegend, Arme seitlich des Pat.	0.098	<.001
Komentieren der Inspektion	-0.1469	<.001
Auskultation in allen 4 Quadranten	-0.0435	.032
Perkussion	-0.185	<.001

Palpation		
Beginn in schmerzloser Region, oberflächliche und tiefe Palpation	-0.0027	.707
Je nach Szenario: Sigmadivertikulitis: Überprüfen des Loslassschmerz Cholezystitis: Murphy Zeichen Appendizitis: Überprüfen des Loslassschmerz (Blumberg-Zeichen),/Mc Burney/ Lanz-Punkt/ Psoas-Zeichen	-0.0218	.195
<i>Verdachtsdiagnose:</i>		
Benennen einer Verdachtsdiagnose Erklären des weiteren Vorgehens (weitere Diagnostik, Therapie)	-0.0476	.136
<i>Umgang mit Patient</i>		
Professionelle Kontaktaufnahme Kommuniziert in Patientengerechter Sprache Zeigt Empathie Hygienisches Arbeiten	-0.0068	.495

Außerdem wurde ein Vergleich der erreichten Punktezahl für alle Items, die der Anamnese oder der körperlichen Untersuchung zugeordnet wurden, durchgeführt.

Es bestand ein signifikanter Unterschied zwischen den erlangten Punkten für die Anamnese zwischen der normalen OSCE Station ohne Telefonanrufe ($M = 2.77$, $SD = 0.76$) und der OSCE Station mit Telefonanrufen ($M = 2.44$, $SD = 0.80$), $t(146) = 4.22$, $p < .001$, $d = 0.36$. Dies deutet darauf hin, dass die Proband*innen mehr Punkte für die Anamnese in der „akutes Abdomen“ Station bekamen, wenn sie nicht durch Telefonanrufe unterbrochen wurden.

Es bestand ebenfalls ein signifikanter Unterschied zwischen den erlangten Punkten für die körperliche Untersuchung zwischen der normalen OSCE Station ohne Telefonanrufe ($M = 3.12$, $SD = 0.68$) und der OSCE Station mit Telefonanrufen ($M = 2.82$, $SD = 0.70$), $t(146) = 4.82$, $p < .001$, $d = 0.40$. Dies deutet ebenso darauf hin, dass die Proband*innen mehr Punkte für die Körperliche Untersuchung in der „akutes Abdomen“ Station bekamen, wenn sie nicht durch Telefonanrufe unterbrochen wurden.

4.2.5. Korrelation zwischen der Leistung in der regulären Station und der Studienstation

Um zu analysieren, ob die Leistung in der regulären Station ohne Telefonanrufe in Zusammenhang mit der Leistung in der Studienstation mit Telefonanrufen stand, wurde eine Korrelationsanalyse zwischen den Punkten aus den jeweiligen Stationen durchgeführt.

In OSCE I war die Korrelation zwischen den Punkten aus der regulären Station und den Punkten aus der Studienstation (mit Telefonanrufen), statistisch signifikant $r(161) = .262$, $p = .001$.

In OSCE II war die Korrelation zwischen den Punkten aus der regulären Station und den Punkten aus der Studienstation (mit Telefonanrufen) ebenfalls statistisch signifikant $r(147) = .348$, $p < .001$.

Dies bedeutet, dass sowohl in OSCE I als auch in OSCE II die Proband*innen, die in der normalen OSCE Station mehr Punkte bekamen, auch in der OSCE Station mit Telefonanrufen mehr Punkte bekamen. Beide Stationen waren demnach geeignet, zwischen leistungsstarken und schwachen Proband*innen zu unterscheiden.

4.3 Korrektes Wiedergeben der Informationen aus dem Telefonanruf

In beiden OSCE-Prüfungen sollten die Proband*innen im zweiten Telefonanruf den Patientennamen und den Blutzuckerwert, welche im ersten Telefonanruf genannt worden waren, wiedergeben. Tabelle 7 zeigt, wie viel Prozent der Proband*innen in OSCE I und II den Namen und den Blutzuckerwert korrekt wiedergeben konnten. Die Tabelle orientiert sich an der Tabelle im vorab veröffentlichten Paper ¹.

Tabelle 7, Anzahl an korrekt wiedergegebenen Informationen aus dem Telefonanruf in OSCE I und OSCE II ¹

	Name	Blutzuckerwert
OSCE I	28% (45 von 161)	46.0% (74 von 161)
OSCE II	32.7% (48 von 147)	57.1% (84 von 147)

Der Unterschied in der Erinnerungsleistung für die Namen zwischen OSCE I und II war nicht signifikant $t < 0.90$.

Der Unterschied in der Erinnerungsleistung für den Blutzuckerwert zwischen OSCE I und II war ebenfalls nicht signifikant, $t < 1.97$. Es zeigte sich jedoch der Trend dahingehend, dass sich die Proband*innen im OSCE II an den Blutzuckerwert eher erinnerten als Proband*innen in OSCE I.

4.3.1. Unterschiede zwischen den Patientennamen

Um zu untersuchen, ob es einen Unterschied in der Erinnerungsleistung zwischen den verschiedenen Patientennamen (Weber, Fischer, Mutlu) gab, wurde eine univariate Varianzanalyse (ANOVA) durchgeführt.

In OSCE I unterschied sich die Erinnerungsleistung an die drei Namen signifikant, $F(2, 157) = 4.22$, $p = .016$, $\eta^2_p = .051$. Die Bonferroni-korrigierte Post-hoc-Analyse ergab, dass der Name Mutlu signifikant schlechter erinnert wurde als der Name Weber ($p=.026$). Alle weiteren Unterschiede waren nicht signifikant ($p>.082$)

In OSCE II gab es ebenfalls einen signifikanten Unterschied in der Erinnerungsleistung zwischen den drei Namen ($F(2, 144) = 6.48$, $p = .002$, $\eta^2_p = .083$). Die Bonferroni-korrigierte Post-hoc-Analyse ergab, dass wie in OSCE I der Name Mutlu signifikant schlechter erinnert wurde als der Name Weber ($p=.001$). Alle weiteren Unterschiede waren nicht signifikant ($p>.246$)

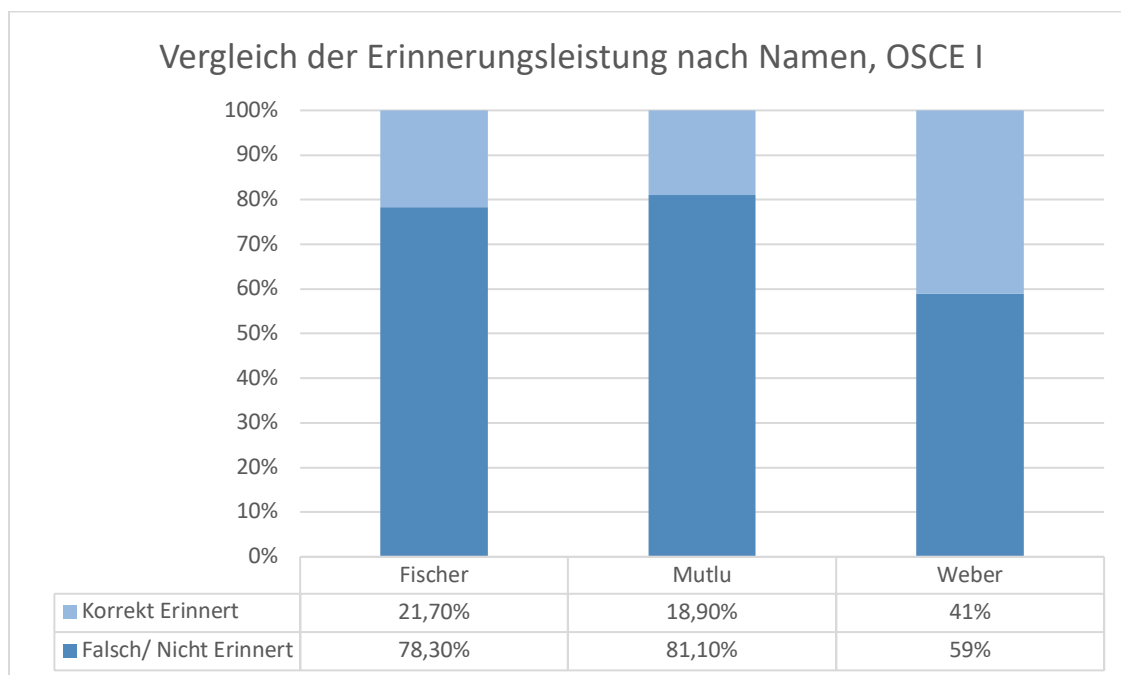


Abbildung 6, Vergleich der Erinnerungsleistung nach Namen, OSCE I

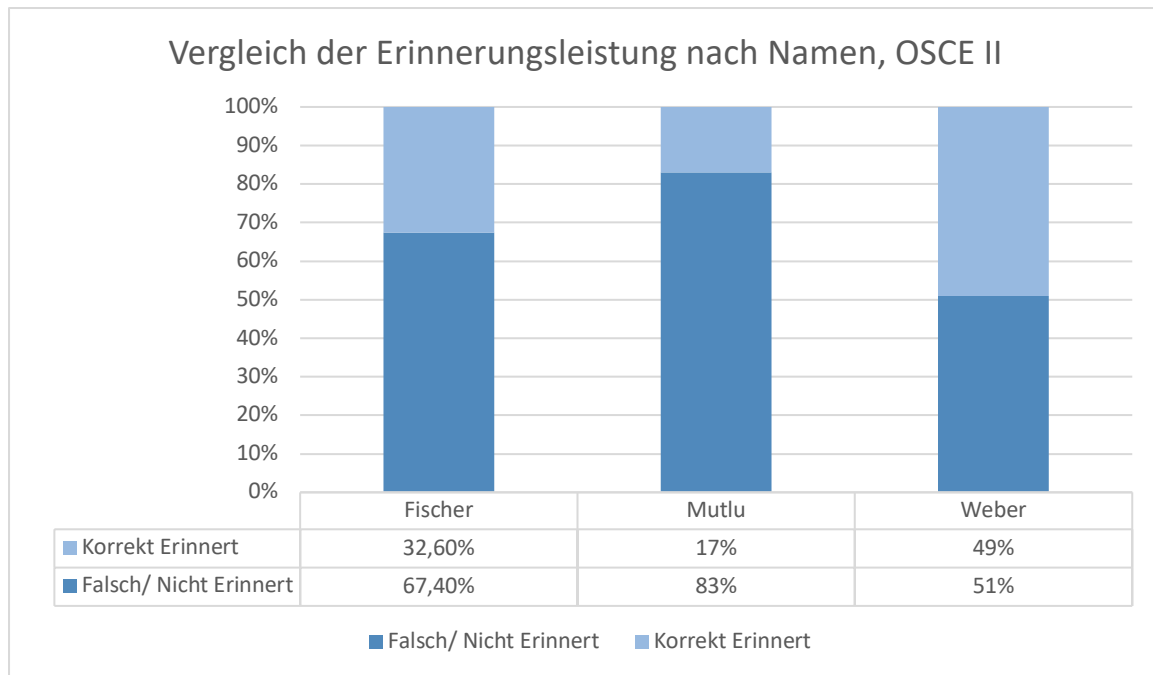


Abbildung 7, Vergleich der Erinnerungsleistung nach Namen, OSCE II

4.4 Prüfer*innen

4.4.1. Unterschiede zwischen den Prüfer*innen der Studie

Um zu untersuchen, ob die verschiedenen Prüfer*innen die Proband*innen unterschiedlich bewerteten, wurde eine univariate Varianzanalyse (ANOVA) durchgeführt.

In OSCE I war in der regulären Station kein signifikanter Unterschied zwischen den Bewertungen der verschiedenen Prüfer*innen, $F(4, 156) = 1.48$, $p = .210$, $\eta^2_p = .037$.

In der Studienstation mit Telefonanrufen in OSCE I gab es ebenfalls keinen signifikanten Unterschied zwischen den Bewertungen der verschiedenen Prüfer*innen, $F(4, 156) = 0.87$, $p = .484$, $\eta^2_p = .022$.

In OSCE II war in der regulären Station kein signifikanter Unterschied in der Bewertung der Proband*innen zwischen den verschiedenen Prüfer*innen, $F(4, 142) = 1.64$, $p = .169$, $\eta^2_p = .044$. Ebenfalls war in der Studienstation in OSCE II kein signifikanter Unterschied in der Bewertung der Proband*innen zwischen den verschiedenen Prüfer*innen, $F(4, 142) = 1.52$, $p = .199$, $\eta^2_p = .041$.

4.4.2. Korrelation zwischen den Prüfer*innen der Studie und den Prüfer*innen der Universität

Die reguläre OSCE Station wurde sowohl durch offizielle Prüfer*innen der Universität bewertet (für die offiziellen Prüfungsergebnisse), als auch durch Prüfer*innen der Arbeitsgruppe (für die Studie). Für die Studie wurden leicht modifizierte Checklisten eingesetzt. Die einzelnen Items stimmten jedoch größtenteils überein und die maximal erreichbare Punktzahl lag wie bei den offiziellen Checklisten bei 10 Punkten.

Um zu untersuchen, ob die Prüfer*innen der Studie (Doktorand*innen) zu ähnlichen Ergebnissen kamen wie die eingesetzten Prüfer*innen der Universität, wurde ein Korrelations-Test durchgeführt.

In OSCE I zeigte sich eine signifikante Korrelation zwischen der Bewertung der eingesetzten Prüfer*innen der Studie und der Bewertung der Prüfer*innen der Universität, $r(161) = .803$, $p < .001$.

In OSCE II bestand ebenfalls eine signifikante Korrelation zwischen der Bewertung der Prüfer*innen der Studie und den Prüfern der Universität, $r(147) = .629$, $p < .001$.

4.5 Ausbildung im medizinischen Bereich

4.5.1. OSCE I

In OSCE I gaben 53 von 161 Proband*innen an, vor dem Studium eine Ausbildung im medizinischen Bereich abgeschlossen zu haben (z.B. als Krankenpfleger*in oder Rettungsassistent*in). In der normalen OSCE Station war ein signifikanter Unterschied in der Punkteanzahl zwischen den Proband*innen mit Ausbildung ($M = 7.16$, $SD = 1.96$) und den Proband*innen ohne Ausbildung ($M = 6.09$, $SD = 1.57$), $t(159) = 3.74$, $p = .001$, $d = 0.60$.

In der OSCE I Station mit Telefonanrufen war ebenfalls ein signifikanter Unterschied in der Punkteanzahl zwischen den Proband*innen mit Ausbildung ($M = 6.61$, $SD = 1.68$) und den Proband*innen ohne Ausbildung ($M = 5.63$, $SD = 1.60$), $t(159) = 3.60$, $p = .001$, $d = 0.58$.

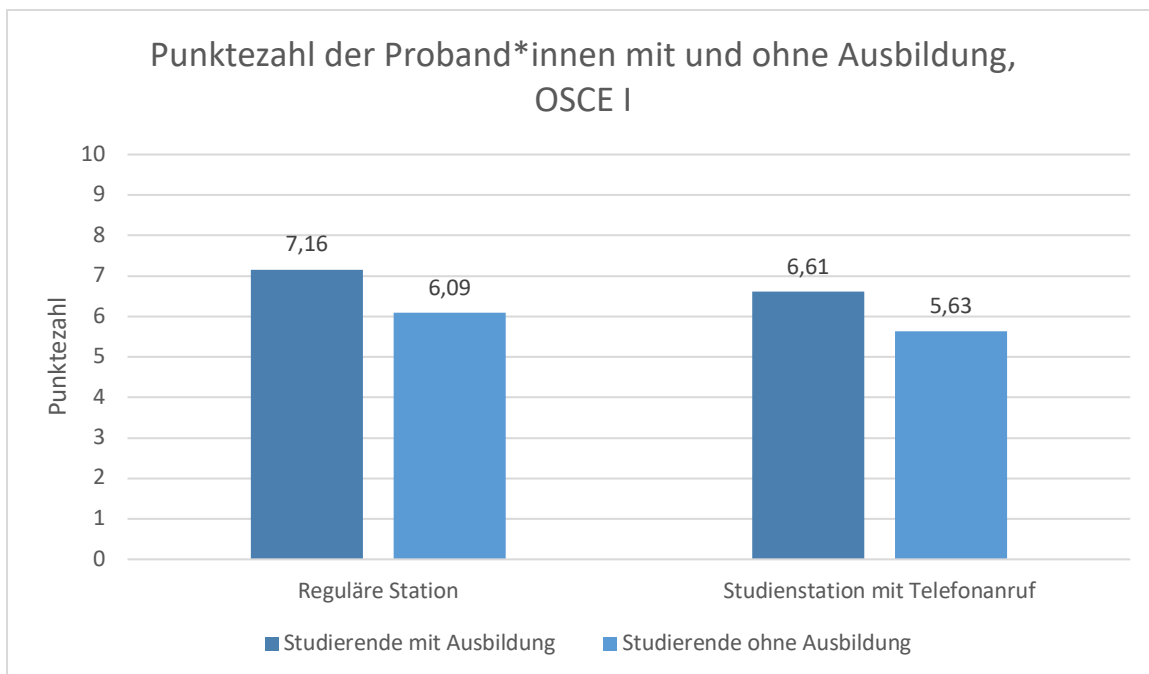


Abbildung 8, Punktezahl der Proband*innen mit und ohne Ausbildung in OSCE I

4.5.2. OSCE II

In OSCE II gaben 27 von 146 Proband*innen an, vor dem Studium eine Ausbildung im medizinischen Bereich abgeschlossen zu haben. Ein Proband der 147 eingeschlossenen Probanden musste aufgrund fehlender Angaben im Fragebogen zum Ausbildungsstand in der Analyse ausgeschlossen werden.

In der normalen OSCE II Station war kein signifikanter Unterschied in der Punkteanzahl zwischen den Proband*innen mit Ausbildung ($M = 7.34$, $SD = 1.46$) und den Proband*innen ohne Ausbildung ($M = 7.65$, $SD = 1.14$), $t(144) = 1.24$, $p = .218$, $d = -0.26$.

In der OSCE II Station mit Telefonanrufen war ebenfalls kein signifikanter Unterschied in der Punkteanzahl zwischen den Proband*innen mit Ausbildung ($M = 6.92$, $SD = 1.17$) und den Proband*innen ohne Ausbildung ($M = 6.82$, $SD = 1.25$), $t(144) = -0.36$, $p = .718$, $d = 0.08$.

4.6 Muttersprache der Studierenden

4.6.1. OSCE I

In OSCE I gaben 15 von 161 Proband*innen an, dass Deutsch nicht ihre Muttersprache ist.

In der regulären OSCE Station gab es keinen signifikanten Unterschied zwischen der Leistung der Proband*innen mit Deutsch als Muttersprache ($M = 6.52$, $SD = 1.76$) und den Proband*innen mit einer anderen Muttersprache, ($M = 5.71$, $SD = 1.75$), $t(159) = 1.69$, $p = .093$, $d = 0.46$.

In der Studienstation mit Telefonanrufen gab es einen signifikanten Unterschied zwischen der Leistung der Proband*innen mit Deutsch als Muttersprache ($M = 6.05$, $SD = 1.68$) und den Proband*innen mit einer anderen Muttersprache ($M = 5.02$, $SD = 1.43$), $t(159) = 2.27$, $p = .025$, $d = 0.66$. Die Proband*innen, die Deutsch nicht als Muttersprache angaben, erreichten in der Studienstation signifikant weniger Punkte.

4.6.2. OSCE II

In OSCE II gaben 24 von 147 Proband*innen im Fragebogen an, dass Deutsch nicht ihre Muttersprache ist.

In der regulären OSCE Station gab es einen signifikanten Unterschied zwischen der Leistung der Proband*innen mit Deutsch als Muttersprache ($M = 7.68$, $SD = 1.19$) und den Proband*innen mit einer anderen Muttersprache, ($M = 7.10$, $SD = 1.21$), $t(145) = 2.19$, $p = .03$, $d = 0.48$. Die Proband*innen mit Deutsch als Muttersprache erreichten in der Station signifikant mehr Punkte.

In der Studienstation mit Telefonanrufen gab es keinen signifikanten Unterschied zwischen der Leistung der Proband*innen mit Deutsch als Muttersprache ($M = 6.88$, $SD = 1.23$) und den Proband*innen mit einer anderen Muttersprache ($M = 6.63$, $SD = 1.25$), $t(145) = 0.93$, $p = .355$, $d = 0.20$.

4.7 Regelmäßige sportliche Betätigung

Um zu untersuchen, ob es einen Zusammenhang zwischen der Leistung der Proband*innen in den OSCE Stationen und einer regelmäßigen sportlichen Betätigung der Proband*innen gab, wurde ein Korrelations-Test zwischen der erreichten Punktezahl in den OSCE Stationen und der von den Proband*innen angegebenen Stundenanzahl Sport pro Woche durchgeführt. Wie Tabelle 8 zeigt, gab es keine signifikante Korrelation zwischen der angegebenen Stundenzahl, die der Proband/ die Probandin pro Woche Sport treiben und der erreichten Punktezahl.

Tabelle 8, Korrelation zwischen sportlicher Betätigung und erreichter Punktzahl

	Korrelation zwischen wöchentlicher Stundenzahl Sport und Punktezahl in der regulären Station	Korrelation zwischen wöchentlicher Stundenzahl Sport und Punktezahl in der Studienstation mit Telefonanrufen
OSCE I	$r(107) = .08, p = .412$	$r(107) = -.01, p = .946$
OSCE II	$r(78) = -.10, p = .399$	$r(78) = .10, p = .364$

4.8 Geschwister

Im Fragebogen gaben die Proband*innen an, wie viele Geschwister sie haben. Um zu analysieren, ob es einen Zusammenhang zwischen der Anzahl an Geschwistern und der erreichten Punktezahl in den OSCE Stationen gab, wurde ein Korrelations-Test durchgeführt. Tabelle 9 zeigt, dass es keine signifikante Korrelation zwischen der Leistung der Proband*innen und der Anzahl an Geschwistern gab.

Tabelle 9, Korrelation zwischen der Geschwisteranzahl und erreichten Punktezahl

	Korrelation zwischen der Geschwisteranzahl und erreichten Punktezahl in der regulären Station	Korrelation zwischen der Geschwisteranzahl und erreichten Punktezahl in der Studienstation mit Telefonanrufen
OSCE I	$r(140) = -.10, p = .266$	$r(140) = .09, p = .293$
OSCE II	$r(130) = .01, p = .937$	$r(130) = .14, p = .116$

4.9 Videospiele

In OSCE I gaben 42 der 161 Proband*innen an, Videospiele zu spielen. Es gab sowohl in der regulären Station als auch in der Studienstation keinen signifikanten Unterschied in der erreichten Punktezahl der Proband*innen, die Videospiele spielten und denen, die keine Videospiele spielen, alle $t < 0.52$.

In OSCE II gaben 36 der 146 Proband*innen an, Videospiele zu spielen. Bei einem Probanden der 147 eingeschlossenen Probanden konnten die Antworten zu Videospiele in dem Fragebogen nicht eingeschlossen werden, da keine Angaben gemacht wurden.

Es gab sowohl in der regulären Station als auch in der Studienstation keinen signifikanten Unterschied in der erreichten Punktezahl der Proband*innen, die Videospiele spielten und denen, die keine Videospiele spielen, alle $t < 1.50$.

4.10 OSCE II Szenarien

Um zu untersuchen, ob die Kombinationen der verschiedenen Szenarien in OSCE II einen Einfluss auf die Leistung der Proband*innen in den Stationen mit Unterbrechung durch Telefonanrufe und ohne Unterbrechung durch Telefonanrufe hatten, wurde eine mixed ANOVA mit einem 6 x 2 Design durchgeführt, mit den Variablen „Szenarien“ und „Unterbrechungen“ (6 Szenarienkombinationen x 2 Stationen, mit Unterbrechungen oder ohne Unterbrechungen). Für die Variable Unterbrechungen wurde die Punktezahl in der regulären Station und die Punktezahl in der Studienstation mit Telefonanrufen betrachtet. Die berücksichtigten sechs Szenarien umfassten alle drei Erkrankungen (Appendizitis, Cholezystitis, Divertikulitis) in den verschiedenen Kombinationen. Da ein Proband/ eine Probandin in der regulären Station und in der Studienstation unterschiedliche Erkrankungen geprüft wurde, ergaben sich unterschiedliche Paarungen (Appendizitis und Cholezystitis, Appendizitis und Divertikulitis sowie Divertikulitis und Cholezystitis). In der Analyse wurde auch berücksichtigt, welche der beiden Erkrankungen die Proband*innen in der Studienstation geprüft wurden. Die Reihenfolge, in welcher die Proband*innen die Szenarien bearbeiteten, wurde nicht berücksichtigt, da diese randomisiert aufgeteilt wurde (56,5% bearbeiteten erst die reguläre Station und dann die Studienstation, 43,5% erst die Studienstation und dann die reguläre Station).

Es trat ein signifikanter Haupteffekt für Unterbrechung auf, was für einen signifikanten Unterschied zwischen den Punktezahlen der Studienstation mit Telefonanrufen und den Punktezahlen in der regulären Station spricht, unabhängig von der durchlaufenen Szenarien Kombination, $F(1, 140) = 51.90.66$, $p < .001$, $\eta^2_p = .270$.

Hinsichtlich der unterschiedlichen Szenarien trat auch ein signifikanter Haupteffekt auf, was für einen signifikanten Unterschied der Punktezahlen zwischen den Szenarien Kombinationen spricht, unabhängig davon, ob die Proband*innen unterbrochen wurden oder nicht. $F(5, 140) = 2.42$, $p = .039$, $\eta^2_p = .079$.

Außerdem gab es einen signifikanten Interaktionseffekt, $F(5, 140) = 2.31$, $p = .047$, $\eta^2_p = .076$. Daraus lässt sich schließen, dass der Effekt der Unterbrechung (gemessen anhand der unterschiedlichen Punktwerte in der Station mit Telefonanrufen und ohne Telefonanrufe) je nach Szenarien Kombination unterschiedlich war.

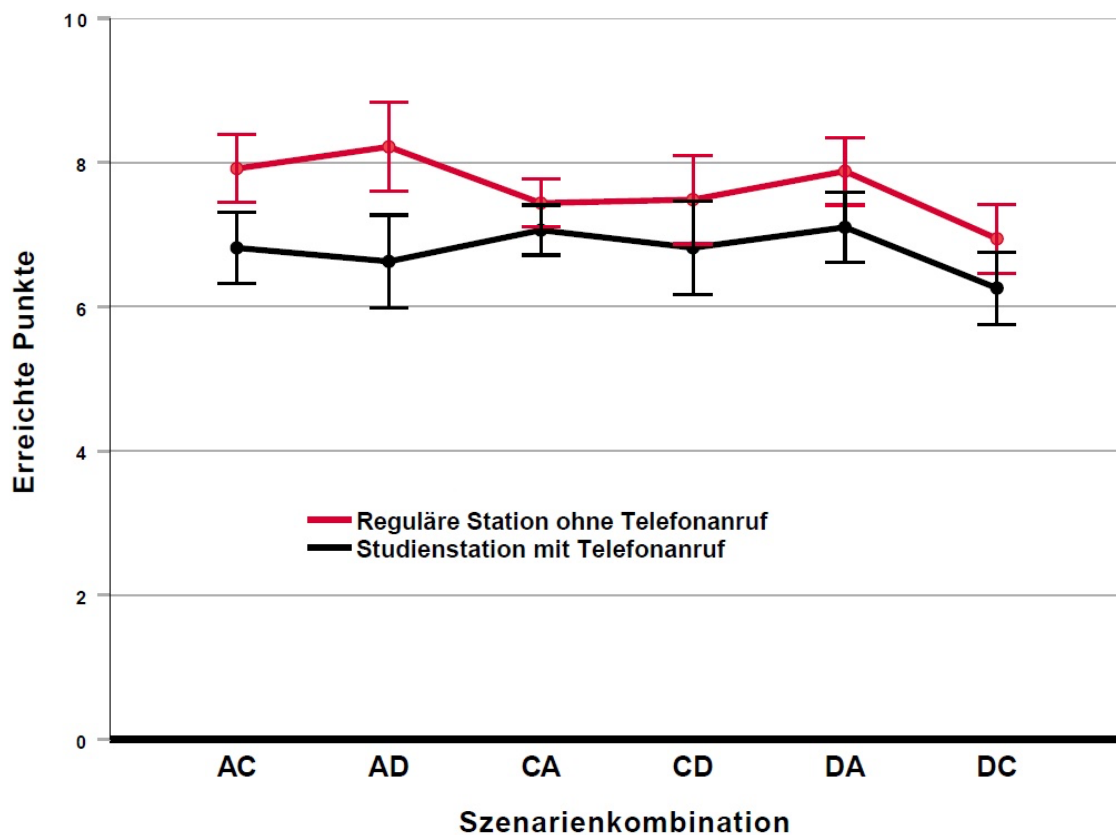


Abbildung 9, Erreichte Punktezahl nach Szenarien Kombination

A: Appendizitis, C: Cholezystitis, D: Divertikulitis. Das jeweils zuerst genannte Szenario in einer Kombination war ohne Telefonanruf, das zweite war das Szenario der Studienstation mit Telefonanruf.

AC: Appendizitis (reguläre Station) und Cholezystitis (Studienstation)

AD: Appendizitis (reguläre Station) und Divertikulitis (Studienstation)

CA: Cholezystitis (reguläre Station) und Appendizitis (Studienstation)

CD: Cholezystitis (reguläre Station) und Divertikulitis (Studienstation)

DA: Divertikulitis (reguläre Station) und Appendizitis (Studienstation)

DC: Divertikulitis (reguläre Station) und Cholezystitis (Studienstation)

Der Abbildung 9 ist zu entnehmen, dass in den Kombinationen AC und AD ein relativ großer Unterschied in der erreichten Punktezahl in der Studienstation und in der regulären Station auftrat. So schnitten die Proband*innen in dem Appendizitis Szenario ohne Unterbrechung besonders gut ab und in dem Cholezystitis, bzw. Divertikulitis Szenario mit Unterbrechung im Vergleich eher schlecht. Eine Überschneidung der Standardabweichung ist nicht zu sehen. In der Kombination CA lagen die Punktezahlen in der Studienstation und regulären Station am nächsten aneinander. Im Vergleich zu den anderen Kombinationen schnitten die Proband*innen im Cholezystitis Szenario ohne Unterbrechung im Vergleich eher schlecht ab und die Studierenden in dem Appendizitis Szenario mit Unterbrechung verhältnismäßig gut. Vergleicht man die durchschnittlich erreichten Punktezahlen in den verschiedenen Szenarien (unabhängig von der Station) zeigt sich der Trend dahingehend, dass in dem Appendizitis Szenario etwas besser abgeschnitten wurde (s.a. Tabelle 10). Insgesamt gab es jedoch keinen signifikanten Unterschied in der Anzahl der erreichten Punkte zwischen den drei Szenarien, $p > .05$.

Tabelle 10, Durchschnittlich erreichte Punktezahl in den verschiedenen Szenarien

Szenario	Appendizitis	Cholezystitis	Divertikulitis
Erreichte Punktzahl (Mittelwert)	7.40	7.06	7.16
SD	1.24	1.33	1.22

In Abbildung 10 ist dargestellt, welche Szenarien Kombinationen wie häufig geprüft wurden. Die verschiedenen Kombinationen wurden zwischen 14 und 48 mal geprüft. Eine gleichmäßige Aufteilung der Paarungen war nicht möglich, da während der Durchführung der OSCE-Prüfung und Datenerhebung Absprachen mit den Prüfer*innen der Universität und den verschiedenen Schauspieler*innen erfolgen mussten. Die Paarungen AC und AD, in welchen der Unterschied zwischen der Studienstation und der regulären Station ausgeprägt auftraten (s.a. Abbildung 8), sind ähnlich häufig wie andere Paarungen geprüft worden. Die Kombination CA, in welcher der Unterschied in den erreichten Punktezahlen nicht so ausgeprägt war (s.a. Abbildung 8), wurde verhältnismäßig oft geprüft (bei 48 von 147 Proband*innen). Die Zuteilung zu den verschiedenen Szenarien Kombinationen erfolgte randomisiert.

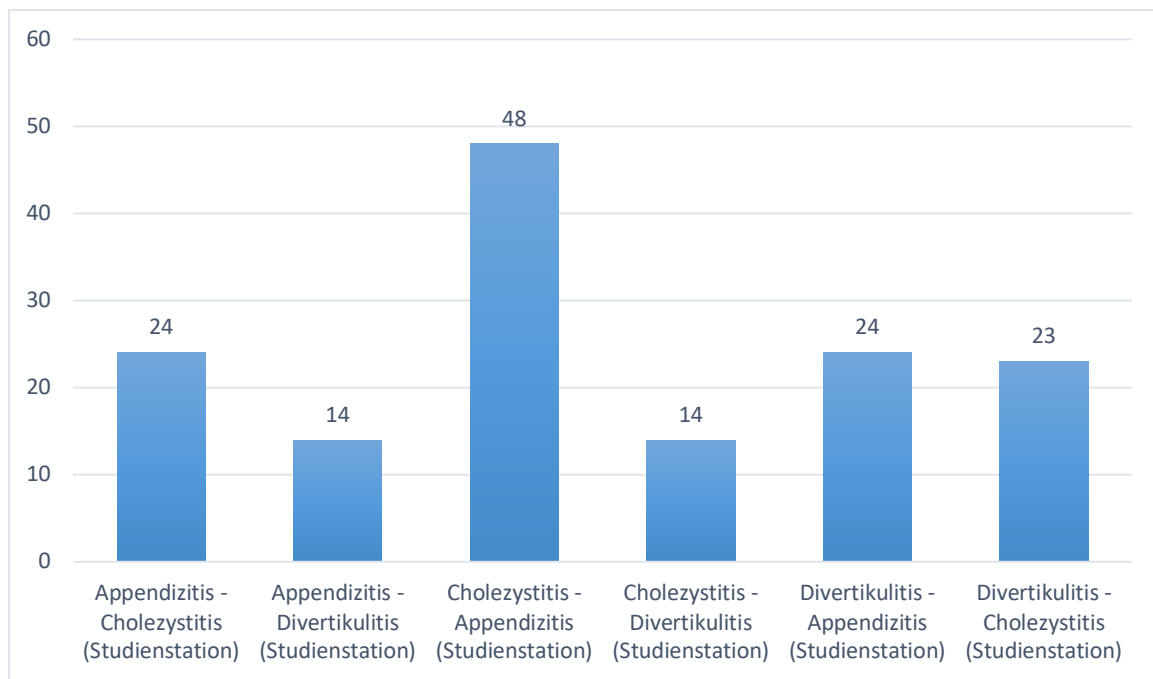


Abbildung 10, Häufigkeitsverteilung der Szenarien Kombinationen

5. Diskussion

5.1 Zusammenfassung der Ergebnisse

Die vorliegende Studie zeigte, dass die Proband*innen sowohl in OSCE I, als auch in OSCE II in der Studienstation mit Telefonanrufen signifikant schlechter abschnitten, als in der Station ohne Unterbrechungen durch Telefonanrufe. Dies war unabhängig davon, ob die Proband*innen rechtzeitig fertig wurden oder nicht.

Außerdem konnten viele Proband*innen die Patientennamen und Blutzuckerwerte, welche im ersten Anruf mitgeteilt wurden, nicht korrekt wiedergeben.

Proband*innen, die vor dem Studium eine Ausbildung im medizinischem Bereich absolvierten schnitten in OSCE I sowohl in der Studienstation, als auch in der regulären Station besser ab. Bei OSCE II bestand kein Unterschied zwischen den Proband*innen mit und ohne medizinische Ausbildung.

Bezüglich weiterer Einflussfaktoren ergab die Ergebnisanalyse, dass die Variablen Sport und Geschwister keinen Einfluss haben. Außerdem zeigte sich, dass nicht alle Proband*innen mit der Aufgabe rechtzeitig fertig wurden und dass es Unterschiede in den Leistungen in den verschiedenen Szenarien in OSCE II gab.

5.2 Diskussion der Ergebnisse: Effekt der Telefonanrufe auf die Leistung der Proband*innen und eine Einordnung in den aktuellen Forschungsstand

Die vorliegende Studie zeigte, dass die Proband*innen sowohl in OSCE I, als auch in OSCE II in der Station mit Telefonanrufen signifikant schlechter abschnitten, als in der Station ohne Unterbrechungen durch Telefonanrufe.

Damit kann die vorliegende Studie nachweisen, dass Telefonanrufe nicht nur im Operationssaal negative Effekte haben, sondern auch bei anderen ärztlichen Tätigkeiten (beim Legen einer Venenverweilkanüle, einer Anamnese und körperlichen Untersuchung) ^{27,28}. Die Ergebnisse entsprechen den Erkenntnissen der experimentellen Studien zu den Effekten von Unterbrechungen durch Telefonanrufe im Operationssaal von Yang et al. und Weigl et al. ^{27,28}. In den Studien machten die Proband*innen in der primären Aufgabe (Operieren) mehr Fehler, wenn sie durch Telefonanrufe abgelenkt wurden ^{27,28}.

Murji et al. und Sujka et al. konnten in ihren Studien hingegen keinen negativen Effekt von Pageranrufen auf die primäre Tätigkeit (Operieren) nachweisen ^{25,26}, allerdings kam es zu Fehlern in der unterbrechenden Tätigkeit. Die Proband*innen beantworteten medizinische Fragen, welche über die Pager gestellt wurden, falsch ^{25,26}.

In experimentellen Studien, die die Effekte von Ablenkungen und Unterbrechungen auf ärztliche Tätigkeiten, die nicht durch Telefone oder Pager bedingt waren, untersuchten, zeigten sich uneinheitliche Ergebnisse.

In einigen Studien wurden signifikant mehr Fehler festgestellt, wenn die Proband*innen abgelenkt oder unterbrochen wurden ^{27,28,80,82,87,89,91,92}, in anderen Studien gab es keine signifikanten Unterschiede ^{25,26,81,83,86,90,93,95,96}.

In einigen Studien brauchten die Proband*innen zudem mehr Zeit für die gleiche Tätigkeit, wenn sie abgelenkt oder unterbrochen wurden, ^{25,82,83,87,90,94–96,121} in anderen war der Unterschied nicht signifikant ^{27,89}. Auch wurde in einigen Studien beobachtet, dass es häufiger zu Falschantworten von Fragen kam, wenn die Proband*innen abgelenkt wurden ^{25,26,81,86,87}, in anderen Studien gab es keinen signifikanten Unterschied ^{88,90}.

Es stellt sich die Frage, warum es in den experimentellen Studien zu so unterschiedlichen Ergebnissen kam. Sind gleichermaßen im klinischen Alltag unterschiedliche Effekte von Unterbrechungen und Ablenkungen von Ärzt*innen zu erwarten und wovon hängen diese unterschiedlichen Effekte ab? Wie können diese unterschiedlichen Effekte im Sinne der Patienten*innensicherheit beeinflusst werden?

In der Ergebnisanalyse der vorliegenden Studie wurden verschiedene Aspekte berücksichtigt, anhand derer die Effekte der Telefonanrufe besser differenziert und Auswirkungen auf die Arbeit von Ärzt*innen abgeleitet werden können. Es erfolgte unter anderem eine Auswertung der verschiedenen Items der Checkliste und der Variablen des Fragebogens sowie eine Auswertung der Bearbeitungszeit.

Im Bereich der kognitiven Psychologie wurden verschiedene Faktoren herausgearbeitet, die für die Auswirkungen von Unterbrechungen entscheidend sein können ^{63,97}. Zu den Faktoren zählen unter anderem der Zeitpunkt der Unterbrechung, die Erfahrungen im Umgang mit Unterbrechung und die Komplexität der primären und sekundären/ unterbrechenden Tätigkeiten ^{63,97}. Diese Faktoren werden im Folgenden in Bezug auf die vorliegenden Studienergebnisse sowie andere Studien diskutiert.

Zudem konnten wie oben beschrieben in experimentellen Studien zu Unterbrechungen und Ablenkungen von medizinischem Personal neben den Effekten auf die primäre Tätigkeit auch Effekte auf die Bearbeitungsdauer einer Tätigkeit und auf die korrekte Ausführung der zweiten, unterbrechenden Tätigkeit (z.B. Falschantworten auf Fragen) festgestellt werden. Dies soll ebenfalls in den folgenden Abschnitten der Diskussion thematisiert werden.

5.2.1. Effekt der Telefonanrufe auf die Bearbeitungsdauer der unterbrochenen Tätigkeit

In OSCE I wurden sowohl in der regulären Station als auch in der Studienstation 75.8% der Proband*innen nicht fertig. Dies spricht für eine Überforderung der Proband*innen mit der Prüfungsaufgabe. Tatsächlich haben viele Proband*innen zum Zeitpunkt der Prüfung bisher nur einen Kurs zum Legen von Venenverweilkanülen belegt, aber wenig praktische Erfahrungen gesammelt, um die Fertigkeiten zu vertiefen. Dies folgt meist erst im Verlauf des Studiums im Rahmen von Famulaturen und dem Praktischen Jahr. Da sowohl in der Studienstation als auch in der regulären Station gleich viele Proband*innen nicht rechtzeitig fertig wurden, ist trotz des hohen prozentualen Anteils von einer Vergleichbarkeit der Ergebnisse auszugehen.

In OSCE II wurden hingegen 14.3% der Proband*innen (21 von 147) in der regulären Station nicht fertig und 23.1% der Proband*innen (34 von 147) in der Station mit Telefonanrufen nicht fertig. Der Unterschied bestand, obwohl den Proband*innen in der Studienstation mehr Zeit für die Telefonanrufe zur Verfügung gestellt wurde. Es lässt sich vermuten, dass die verlängerte Zeit nicht ausreichte, die Unterbrechungen wieder auszugleichen. Möglicherweise waren die Proband*innen durch die Unterbrechung insgesamt langsamer in der Bearbeitung der Prüfungsaufgabe in OSCE II.

Der Einfluss von Unterbrechungen auf die Bearbeitungsdauer der unterbrochenen Aufgabe wurde in verschiedenen Studien untersucht.

So wurde in Beobachtungsstudien festgestellt, dass Unterbrechungen und Ablenkungen zu längeren OP Zeiten führen können ^{76,78}. In experimentellen Studien im OP konnte ebenfalls eine längere OP Dauer bei Ablenkungen und Unterbrechungen festgestellt werden ^{25,82,83,87,88,90,94}.

Auch eine Diagnosestellung brauchte in experimentellen Studien durch eine Unterbrechung bzw. das gleichzeitige Ausführen von zwei Tätigkeiten länger ^{95,96}.

Wie es zu einer verlängerten Bearbeitung von Aufgaben durch Ablenkungen und Unterbrechungen kommen kann, kann anhand verschiedener Modelle der kognitiven Psychologie erklärt werden. In dem Memory for Goals Modell wird davon ausgegangen, dass nach einer Unterbrechung das kognitive Ziel der vorherigen Tätigkeit wieder aktiviert werden muss ^{62,122–125}. Dies kann einige Zeit dauern und damit zu einer Verzögerung führen, bis die initiale Tätigkeit tatsächlich wieder aufgenommen wird (sogenannte Resumption Lag) ^{62,122–125}. Einfluss auf diese Verzögerung kann auch die Arbeitsgedächtnis Kapazität haben ¹²⁶.

Eine Überforderung der Arbeitsgedächtnis Kapazität kann z.B. durch eine große Anzahl an Informationen zustande kommen ^{127,128}.

Bezieht man dies auf die vorliegende Studie, könnte die längere Bearbeitungsdauer der Anamnese- und körperlichen Untersuchung in OSCE II in der Studienstation auf eine größere Anzahl an Informationen, welche von den Proband*innen verarbeitet werden muss, zurückgeführt werden.

Ein anderer Erklärungsansatz aus der kognitiven Psychologie bezieht sich auf das Paradigma Task-switching (Aufgabenwechsel). Es wird davon ausgegangen, dass für jede Tätigkeit/kognitive Leistung ein entsprechendes kognitives Schema oder Task-set benötigt wird, um ein Ziel zu erreichen ^{129,130}. Bei einem Task-switch, also dem Wechsel einer Tätigkeit, muss ein neues Task-set ausgewählt und rekonfiguriert werden ^{129,130}. Als Switch-Cost wird das Phänomen bezeichnet, dass es bei einem Wechsel zwischen verschiedenen Tätigkeiten zu Verzögerung und zu Fehlern kommen kann ^{129,130}. Rogers und Monsell führen dies vor allem auf den Prozess der endogen kontrollierten Task-Set Rekonfiguration zurück ¹³⁰.

Allport et al. hingegen gehen in ihrem Task Set Inertia Modell davon aus, dass eine Unterdrückung von Task-Sets für Switch-Cost verantwortlich ist ^{131,132}. Wenn von einer Tätigkeit zu einer anderen konkurrierenden Tätigkeit gewechselt werden soll, muss das Task-set der aktuellen Tätigkeit unterdrückt werden (negatives Priming) und das Task-Set der neuen Tätigkeit wird durch positives Priming verstärkt ^{131,132}. Wenn dann die vorherige Tätigkeit wieder aufgenommen werden soll, ist dies u.U. nur verzögert möglich, weil das entsprechende Task Set nach wie vor unterdrückt wird ^{131,132}. In Bezug auf die vorliegende Studie ist davon auszugehen, dass für ein Patientengespräch oder eine körperliche Untersuchung ein anderes Task-set benötigt wird als für ein Telefongespräch und es daher zu einem Task-switch durch die Unterbrechung und dadurch zu einer Verzögerung kam.

Unterbrechungen durch Telefonanrufe können demnach nicht nur einen negativen Effekt auf die primäre Tätigkeit haben, sondern auch zu Verzögerungen im Ablauf im klinischen Alltag führen. Dadurch bleibt unter Umständen weniger Zeit für andere Aufgaben und die Entstehung von Zeitdruck und Stress könnten begünstigt werden. Ein erhöhtes Stresslevel ⁹⁻¹⁶ als Folge von Unterbrechungen und Ablenkungen wurde bereits in vielen Studien nachgewiesen.

5.2.2. Effekt der Telefonanrufe auf Teilaufgaben der Prüfung und einzelne Arbeitsschritte

Es zeigte sich, dass die Proband*innen in der Studienstation mit Telefonanrufen in OSCE I signifikant weniger Punkte für das Richten der Materialien, Handschuhe anziehen, Anlegen des Stauschlauchs und zügiges Punktieren erhielten. Dies könnte auf die zuvor ermittelten Zeitpunkte, zu denen die Anrufe erfolgen sollten, zurückgeführt werden. Der erste Anruf sollte während dem Richten der Materialien erfolgen, der zweite kurz vor, nach oder während der Punktion. In beiden Bereichen schnitten die Proband*innen signifikant schlechter ab.

Ebenfalls erreichten die Proband*innen in der Studienstation mit Telefonanrufen in OSCE II signifikant weniger Punkte für einen Teil der Anamnese (Vorerkrankungen und -operationen sowie Allergien und Medikamente) und die körperliche Untersuchung (Inspektion, Auskultation und Perkussion). Auch diese beiden Aufgabenbereiche wurden jeweils durch einen Anruf unterbrochen. Nur für die korrekte Lagerung des Patienten/ der Patientin erhielten die Studierenden in OSCE II in der Studienstation signifikant mehr Punkte als in der regulären Station. Die Lagerung des Patienten/ der Patientin erfolgte meist zu Beginn der Prüfung, also bevor die Unterbrechungen auftraten. Daher war ein negativer Effekt in der Studienstation nicht zu erwarten. Warum es zu einem positiven Effekt kam, bleibt unklar. Möglicherweise waren die Proband*innen zu Beginn in der Studienstation besonders konzentriert und dachten signifikant häufiger daran, eine korrekte Lagerung durchzuführen.

Für andere Unterpunkte war der Unterschied in OSCE I und OSCE II jedoch nicht signifikant. Unter anderem bei „Benennen einer Verdachtsdiagnose und erklären des weiteren Vorgehens (weitere Diagnostik, Therapie)“ war der Unterschied in der erreichten Punktezahl in OSCE II nicht signifikant. Dies ist im Sinne der Patient*innensicherheit zunächst positiv zu werten. Auch in den experimentellen Studien von Monteiro et al. und Cao et al. konnte keine Reduktion in der Anzahl der korrekt gestellten Diagnosen festgestellt werden, wenn die Proband*innen abgelenkt wurden, bzw. zwei Aufgaben gleichzeitig bearbeiten mussten^{95,96}. Das besondere an der vorliegenden Studie ist jedoch, dass der gesamte Ablauf der Diagnosefindung mit einer Anamneseerhebung und einer körperlichen Untersuchung geprüft und umfassend bewertet wurde. Dadurch konnte zwar kein Effekt auf die korrekte Diagnosestellung nachgewiesen werden, aber andere relevante Defizite, beispielsweise bezüglich der Medikamentenanamnese, wurden herausgestellt. Damit prüfte die vorliegende Studie wesentlich umfangreicher den Effekt von Unterbrechungen und Ablenkungen und konnte mögliche negative Konsequenzen nachweisen, die in den Studien von Monteiro et al. und Cao et al. nicht aufgezeigt werden konnten.

5.2.3. Zeitpunkt der Unterbrechung als relevanter Faktor für den Effekt der Unterbrechung

In verschiedenen Reviews wird der Zeitpunkt der Unterbrechung als relevanter Faktor für den Effekt der Unterbrechung identifiziert ^{63,97}.

Eine Unterbrechung soll besonders stören, wenn sie während der Durchführung einer Tätigkeit/ Aufgabe oder während dem Zwischenschritt einer Tätigkeit auftritt und nicht zwischen zwei unterschiedlichen Tätigkeiten ^{63,97}. Das Vergessen bestimmter Schritte einer Tätigkeit tritt außerdem am ehesten direkt nach einer Unterbrechung auf. Dies passiere beispielsweise, wenn in Anlehnung an das Memory for Goals Modell keine Hinweisreize für eine erneute assoziative Verknüpfung nach der Unterbrechung zur Verfügung stehen ^{63,97}.

Diese Hinweisreize können helfen, ein Ziel (mentale Intention eine bestimmte Tätigkeit oder Aufgabe durchzuführen) wieder zu aktivieren, wenn es zuvor aufgrund einer Unterbrechung ausgesetzt wurde ^{63,97}.

Zudem sei eine Unterbrechung vor allem in der Mitte und am Ende einer Tätigkeit störend, da dann nach der Unterbrechung eine größere Menge von Informationen wieder reaktiviert werden müsse ⁶³. Da die Anrufe in der Studie in der Mitte der Prüfung getätigt wurden und zu Zeitpunkten, als die Teilschritte „Richten der Materialien“ und „Punktion“ in OSCE I, bzw. „Anamnese“ und „körperliche Untersuchung“ in OSCE II jeweils noch nicht abgeschlossen waren, führten diese in der vorliegenden Studie zu einem entsprechenden negativen Effekt. Tatsächlich stellten Westbrook et al. in ihrer Beobachtungsstudie fest, dass in 18,5% der Fälle eine unterbrochene Aufgabe vorzeitig beendet und nicht wieder aufgenommen wurde, nachdem sich ein Arzt oder eine Ärztin mit einer Unterbrechung befasst hatte ¹³³. Möglicherweise schlossen die Proband*innen in der vorliegenden Studie durch die Telefonanrufe verschiedener Einzelschritte in der Bearbeitung der Prüfungsaufgabe nicht ab und bekamen dadurch weniger Punkte in der Studienstation.

Diese Erkenntnisse sollten in der Aufklärung medizinischen Personals über mögliche negative Effekte durch Unterbrechungen durch Telefonanrufe berücksichtigt werden und entsprechende Handlungsempfehlungen ausgesprochen werden. Beispielsweise sollte ein Anruf wenn möglich nicht in der Mitte oder kurz vor Ende einer Tätigkeit angenommen werden, weil es dann am ehesten zu Fehlern und zu dem Vergessen von Arbeitsschritten kommen kann. Alternativ kann erwogen werden einen Anruf entgegenzunehmen, wenn zumindest ein Teilschritt einer Aufgabe abgeschlossen wurde und bevor der nächste Teilschritt begonnen wird (z.B. nach dem Abschluss der Anamnese und vor dem Beginn der körperlichen Untersuchung).

5.2.4. Effekt der Telefonanrufe auf die zweite Tätigkeit und auf Gedächtnisleistungen: Korrektes Wiedergeben der Informationen aus den Telefonanrufen

Ein großer Anteil der Proband*innen konnte weder den Patientennamen noch die Blutzuckerwerte korrekt wiedergeben. Allerdings ist den Proband*innen vorher nicht mitgeteilt worden, inwiefern der Inhalt der Anrufe eine Relevanz für die Prüfung habe. Viele haben sich möglicherweise mehr auf die primäre Aufgabe (welche der offiziellen Prüfungsaufgabe entsprach) konzentriert und weniger auf die Anrufe. Zudem hat medizinisches Personal in der Realität die Möglichkeit, sich Informationen, die ihnen beispielweise während eines Telefonanrufs genannt werden, aufzuschreiben. Dies war in der Prüfungssimulation nicht möglich.

Dennoch gibt es einige Studien, in denen die Gedächtnisleistungen bei Unterbrechungen, Ablenkungen und Multitasking ebenfalls schlechter waren.

So sollten in der experimentellen Studie von Cao et al. Medizinstudierende unter verschiedenen Bedingungen eine abstrakte diagnostische Entscheidung treffen und Gedächtnisaufgaben bearbeiten. Wenn sie beides parallel als Dual-Task durchführten, machten die Proband*innen signifikant mehr Fehler bei der Gedächtnisaufgabe, als wenn sie diese als Single-Task bearbeiteten ⁹⁵.

Auch in der Studie von Feuerbacher et al. hatten Ablenkungen und Unterbrechungen einen Einfluss auf die Gedächtnisleistung. So vergaßen die Proband*innen signifikant häufiger zu einem vorher definierten Zeitpunkt, einen Zwischenstand zum Operationsverlauf an das OP-Team zu kommunizieren, wenn sie während der simulierten Operation abgelenkt oder unterbrochen wurden ⁸⁰.

In den Studien von Murji et al. und Sujka et al. wurden Proband*innen während einer simulierten Operation durch Pager unterbrochen und sollten zu anderen Patientenfällen Fragen inhaltlich korrekt beantworten. In beiden Studien hatten die Ablenkungen keinen Einfluss auf die OP-Leistungen, allerdings beantworteten viele der Proband*innen die Fragen falsch oder mit sicherheitskritischen Konsequenzen für die Patient*innen ^{25,26}.

In der Studie von Woods et al. erreichten die Proband*innen weniger Punkte für das korrekte Beantworten von Fragen zu einem Anamnesegespräch, wenn sie dabei eine andere Tätigkeit gleichzeitig ausführten ¹²¹. Bei der zweiten Tätigkeit handelte es sich um das Legen eines intravenösen Zugangs. Interessanterweise konnte bei der gleichzeitigen Bearbeitung der Aufgaben jedoch kein Effekt auf das korrekte Legen des Zugangs beobachtet werden ¹²¹. Die Autor*innen vermuteten, dass die Kombination aus beiden Aufgaben zu einer Überlastung führte und die Proband*innen sich daher auf die für sie schwierigere und fordernde Aufgabe, das Zugang legen, konzentrierten und deshalb die zweite Aufgabe (einer Anamnese zuhören, Fakten erinnern und anschließend wiedergeben) vernachlässigten ¹²¹.

Im Gegensatz dazu konnte die vorliegende Studie einen negativen Einfluss durch Ablenkung durch Telefonanrufe auf das korrekte Legen eines venösen Zugangs nachweisen. Folgt man der oben genannten Annahme von Woods et al. gelang es den Proband*innen in der vorliegenden Studie möglicherweise nicht, sich ausreichend auf die Aufgabe zu fokussieren. Es kann auch angenommen werden, dass die Telefonanrufe eine stärkere Ablenkung generierten als die Fragen zu einem Anamnesegegespräch.

Berücksichtigt werden sollte auch die Methodik, mit welcher das Legen eines venösen Zugangs bewertet wurde. Auch Woods et al. nutzen eine Checkliste, die der hier verwendeten ähnelt, sie ist jedoch weniger ausführlich und fasst mehr Unterpunkte zusammen.

Woods et al. veröffentlichten jedoch keine Analyse der Punktezahl der einzelnen Items. Es wäre beispielsweise interessant, ob es hier zwar keinen Unterschied in der Gesamtpunktzahl gab, aber möglicherweise signifikante Unterschiede in einzelnen Arbeitsschritten, wie es in der vorliegenden Studie nachgewiesen wurde (z.B. Handschuhe anziehen).

Eine mögliche Erklärung für negative Effekte, die auftreten können, wenn zwei Aufgaben gleichzeitig bearbeitet werden, stellte Wickens mit seiner multiple resource theory auf. Diese besagt, dass es zu Störungen kommen kann, wenn zwei parallel ausgeführte Aufgaben ähnliche kognitive Ressourcen benötigen, welche limitiert sind ^{134,135}. Auch in der vorliegenden Studie mussten sich die Proband*innen auf zwei Aufgaben gleichzeitig konzentrieren- das Legen eines venösen Zugangs (in OSCE I) oder ein Anamnesegegespräch und eine körperliche Untersuchung (in OSCE II) und die Telefonanrufe, bzw. das Erinnern an Informationen aus dem ersten Telefonanruf.

Allerdings gab es in den meisten oben genannten Studien eine Kontrollgruppe, in der die Leistung der Proband*innen der zweiten, unterbrechenden Tätigkeit als Single-Task getestet wurde, sodass ein Vergleich der Leistungen unter Single- und Dual-Task Bedingungen möglich war.

Dies erfolgte in der vorliegenden Studie nicht, sodass sich aus den Ergebnissen nicht ableiten lässt, ob die Erinnerungsleistung der Proband*innen an die Patientendaten aus den Telefonanrufen besser gewesen wäre, wenn sie nicht während einer Tätigkeit unterbrochen worden wären. Zu dem Zweck wäre ein anderes Studiendesign nötig gewesen.

Die meisten Studien zu Multitasking und Unterbrechungen von medizinischem Personal konzentrierten sich bisher auf den Effekt von Unterbrechungen und Multitasking auf die primäre Tätigkeit. Die Ergebnisse der vorliegenden Studie und einiger anderer Studien zeigten jedoch auch negative Effekte für die zweite Tätigkeit.

Dies kann unter Umständen zu sicherheitsrelevanten Konsequenzen führen, so wurden in der Studie von Murji et al. klinische Entscheidungen getroffen, die als potenziell gefährlich eingestuft wurden ²⁵. Es ist auch vorstellbar, dass das Vergessen von Patientennamen oder Laborwerten, wie es in unserer Studie festgestellt wurde, zu Fehlern und Verwechslungen führen könnte. Eine weitere wissenschaftliche Auseinandersetzung mit den Effekten von Multitasking und Unterbrechungen auf die zweite Tätigkeit scheint daher sinnvoll. Zudem sollte in der Schulung von medizinischem Personal auf die möglichen negativen Effekte von Unterbrechungen und Multitasking auf alle Aufgaben eingegangen werden. Dies ist sowohl für die Person relevant, die einen Anruf entgegennimmt und beispielsweise während des Telefonats sicherheitskritische Informationen erhält oder Entscheidungen trifft, als auch für die Person, die den Anruf tätigt.

5.2.5. Effekt der Telefonanrufe abhängig von Ausbildung und Berufserfahrung: Erfahrungen im Umgang mit Unterbrechung als relevanter Faktor für den Effekt von Unterbrechungen

In OSCE I schnitten die Proband*innen mit einer Ausbildung im medizinischen Bereich signifikant besser ab als die Proband*innen ohne Ausbildung, sowohl in der Studienstation als auch in der regulären Station. In OSCE II hingegen bestand kein Unterschied zwischen den erreichten Leistungen der Proband*innen mit medizinischer Ausbildung und ohne medizinische Ausbildung. Dies könnte an der Prüfungsaufgabe in OSCE I und OSCE II liegen. Das Legen einer Venenverweilkanüle, wie es in OSCE I geprüft wurde, ist Teil der Tätigkeiten von medizinischen Ausbildungsberufen wie Rettungsassistent*innen. Dadurch hatten die Proband*innen mit medizinischer Ausbildung einen Vorteil.

Außerdem ist davon auszugehen, dass zum Zeitpunkt von OSCE I im fünften Studiensemester der Erfahrungsvorsprung der Proband*innen mit Berufserfahrung noch stärker ausgeprägt war, dieser sich aber im Verlauf des Studiums reduzierte. Im neunten Semester, zum Zeitpunkt der OSCE II Prüfung, hatten die Proband*innen ohne Ausbildung bereits mehr Praktika und damit Möglichkeiten, mehr Erfahrungen zu sammeln.

Allerdings konnte kein Effekt der Unterbrechungen auf die erreichte Punktezahl, abhängig von einer vorherigen medizinischen Ausbildung festgestellt werden.

Andere experimentelle Studien differenzierten ebenfalls zwischen Proband*innen mit Arbeitserfahrung und konnten anders als in dieser Studie unterschiedliche Effekte von Unterbrechungen und Ablenkungen herausarbeiten.

In der Studie von Hsu et al. konnten die unerfahrenen Proband*innen eine laparoskopische Aufgabe trotz Ablenkung genauso gut durchführen wie Proband*innen mit mehr Erfahrung, aber sie konnten insgesamt signifikant weniger Fragen im gleichen Zeitfenster beantworten und die Anzahl der richtig beantworteten Fragen war signifikant geringer ⁸¹.

In einer ähnlichen Studie von Ghazanfar et al. brauchten die unerfahrenen Proband*innen bei der Ablenkung durch eine schwerere Aufgabe signifikant länger, um eine laparoskopische Aufgabe zu bearbeiten ⁸⁷. Außerdem benötigten sie mehr Bewegungen und machten signifikant mehr Fehler. Auf die chirurgische Leistung der erfahrenen Proband*innen hatte die Ablenkung keinen Einfluss ⁸⁷.

Auch im Review von Li et al. wurden Erfahrungen und Geübtheit im Umgang mit Unterbrechungen als ein Faktor identifiziert, welcher Einfluss auf den Effekt von Unterbrechungen habe. Zum einen könne man besser mit Unterbrechungen umgehen, wenn man in der unterbrochenen Tätigkeit routiniert und geübt sei ⁹⁷. Zum anderen könne eine

Gewöhnung an Unterbrechungen und ein geübter Umgang zu einer geringeren Beeinträchtigung der primären Tätigkeit führen⁹⁷.

Unter anderem falle dann im Sinne des activation-based Memory for Goals Modell eine Nutzung von Cues/ Hinweisreizen und dadurch eine Assoziation mit den verschiedenen Schritten der primären unterbrochenen Tätigkeit leichter⁹⁷.

Ein weiterer relevanter Faktor nach Li et al. ist die Strategie, wie mit der Unterbrechung umgegangen wird⁹⁷.

Es ist zu vermuten, dass mit zunehmender Berufserfahrung nicht nur das technische Können und Wissen steigt, sondern auch die Fähigkeit, eine Situation richtig einzuschätzen und die bestmögliche Strategie im Umgang mit Unterbrechungen zu wählen. Li et al. nennen als mögliche Strategien bei einer Unterbrechung durch Telefonanrufe das sofortige Annehmen des Anrufs oder ein kurzes Innehalten vor dem Gespräch, um sich auf die gerade durchgeführte Tätigkeit zu besinnen, oder ein Verschieben des Telefongesprächs auf einen späteren Zeitpunkt⁹⁷. Auch könne die primäre Tätigkeit während des Telefongesprächs gänzlich pausiert werden (Task-switch) oder parallel fortgeführt werden (Multitasking, Dual-Task)⁹⁷.

Zusammenfassend lässt sich feststellen, dass Arbeitserfahrung ein wichtiger Faktor dafür sein kann, welche Effekte Unterbrechungen und Ablenkungen auf die Arbeit von medizinischem Personal haben können. Das Risiko für negative Effekte ist geringer, wenn die unterbrochene Person über mehr Erfahrung im Umgang mit Unterbrechungen und Ablenkungen verfügt.

Demnach sollte der Sensibilisierung für mögliche negative Effekte und einer Berücksichtigung in der Lehre vor allem bei Berufsanfänger*innen und Assistenzärzt*innen eine höhere Relevanz zugesprochen werden.

5.2.6. Effekt der Telefonanrufe abhängig von der Komplexität der primären und sekundären, unterbrechenden Tätigkeiten (Unterschiede zwischen den OSCE II Szenarien)

In OSCE II wurden drei unterschiedliche Szenarien geprüft. Der Effekt der Unterbrechungen (gemessen anhand der unterschiedlichen Punktwerte in der Station mit Telefonanrufen und ohne Telefonanrufe) war je nach Szenarien Kombination unterschiedlich.

Der unterschiedlich stark ausgeprägte Effekt der Unterbrechungen kann möglicherweise auf unterschiedliche Anforderungsprofile der Szenarien zurückgeführt werden.

Dies passt zu den Erkenntnissen anderer Studien, die besagen, dass der mögliche Effekt der Ablenkung oder Unterbrechung vom Anforderungsprofil der primären Tätigkeit („Komplexität der Tätigkeit“) abhängen kann ⁶³. Bei einer komplexen primären Tätigkeit kann es eher zu negativen Effekten kommen, dies erklären sich einige Autor*innen mit einer längeren Codierungs-Phase ⁶³. In dieser Phase wird das Ziel und das Task Set der primären Tätigkeit für eine spätere Wiederaufnahme kodiert ⁶³. Wenn die primäre Tätigkeit komplexer ist, dauere die Codierungs-Phase länger und es kann eher zu Fehlern kommen ⁶³.

Außerdem könne es bei einem Task-switch zu Störungen und Übertragungseffekten vom kognitiven Schema der primären Tätigkeit kommen ⁶³.

Yang et al. untersuchten in ihrer Studie den Einfluss der Komplexität der Tätigkeiten und konnten die oben genannte Annahme bestätigen ²⁷. In dem Studiendesign berücksichtigten die Autor*innen sowohl die Komplexität der primären Tätigkeit (Laparoskopie Aufgabe) als auch die Komplexität der zweiten unterbrechenden Tätigkeit (Telefonanrufe) ²⁷. Sie konzipierten in ihrer Studie jeweils eine leichte und eine komplexe Aufgabe und eine leichte und eine starke Unterbrechung/ Ablenkung ²⁷. Bei der leichteren Aufgabe (Transfer von Objekten) wurde nur bei starker Ablenkung ein Einfluss auf die Leistung der Proband*innen beobachtet. Bei der als schwieriger erachteten Aufgabe (präzises Schneiden) führten beide Formen der Ablenkung zu einer schlechteren Leistung ²⁷. Als leichte Ablenkung wurde nur ein Telefonanruf genutzt, als schwere Ablenkung zwei aufeinander folgende Anrufe mit Rückfragen im zweiten Anruf zu Gesprächsinhalten aus dem ersten Telefongespräch ²⁷.

Ghazanfar et al. nutzen in ihrer Studie ebenfalls zwei unterschiedlich starke Ablenkungen. Sie konnten ebenfalls nur bei der starken Ablenkung einen signifikanten Unterschied in der Fehleranzahl in einer chirurgischen Aufgabe nachweisen ⁸⁷.

Es scheint demnach nicht nur die Komplexität der primären Tätigkeit, sondern auch die Komplexität der sekundären, unterbrechenden Tätigkeit entscheidend für die Effekte zu sein. In Anlehnung an die Studie von Yang et al. wurde in der vorliegenden Studie eine schwere Ablenkung, ebenfalls mit zwei Telefonanrufen, genutzt.

Auch nach Couffe et al. ist neben der Komplexität der primären Tätigkeit die Komplexität der unterbrechenden Tätigkeit für die erfolgreiche Wiederaufnahme der primären Tätigkeit relevant ⁶³. Dies kann mit der Arbeitsgedächtnis Kapazität erklärt werden. Wenn es sich um eine komplexe Unterbrechung handelt und diese verarbeitet werden muss, kann die Aufrechterhaltung der Informationen der primären Tätigkeit beeinträchtigt sein und dies zu einer geringeren Aktivierung des kognitiven Schemas der primären Tätigkeit führen. Dazu kann es kommen, wenn die unterbrechende Tätigkeit sehr komplex ist und einen großen Anteil der Arbeitsgedächtnis Kapazität in Anspruch nimmt ^{63,127,128,136}.

Es ist davon auszugehen, dass für die Proband*innen in der vorliegenden Studie die Telefonanrufe einer schweren Unterbrechung mit komplexer Anforderung entsprachen und diese dadurch einen nachweisbaren negativen Effekt auf die primäre Tätigkeit hatten.

Entsprechende Folgestudien bezüglich des Einflusses der Komplexität von Tätigkeiten könnten interessant sein und im Interesse der Forschung zu Patient*innensicherheit. So hat es auch eine praktische Relevanz zu ergründen, bei welchen primären Tätigkeiten und welchen sekundären Tätigkeiten, bzw. Unterbrechungen ein höheres Risiko für negative Effekte zu erwarten sind und bei welchen ein geringeres Risiko besteht. Die Aufgaben mit einem erhöhten Risiko für negative Effekte, welche insbesondere als sicherheitskritisch einzuschätzen sind, könnten entsprechend geschützt werden.

Beispielsweise gibt es im Bereich der Pflege Ansätze, Störungen beim Richten von Medikamenten zu reduzieren ^{137,138}. Wie aus Studien bekannt treten hierbei besonders häufig Fehler durch Ablenkungen und Unterbrechungen auf und Medikationsfehler können die Patient*innensicherheit in besonderem Maße gefährden ^{139,140}.

5.2.7. Effekt der Telefonanrufe abhängig von der Variable Muttersprache

Um auszuschließen, dass Unterschiede in den Leistungen in den OSCE Stationen aufgrund von Sprachbarrieren bestanden, wurden in den Fragebögen die Muttersprache der Proband*innen erfasst.

In OSCE I gab es in der regulären Station keinen signifikanten Unterschied zwischen den Proband*innen mit Deutsch als Muttersprache und den Proband*innen mit einer anderen Muttersprache. In der Studienstation gab es hingegen einen signifikanten Unterschied. Im Mittel erreichten die Proband*innen mit einer anderen Muttersprache 5.02 Punkte, die Proband*innen mit Deutsch als Muttersprache 6.05 von 10 möglichen Punkten.

Dies führt zu der Hypothese, dass Verständnisprobleme während der Telefongespräche zu einer stärkeren Störung der Proband*innen und damit zu einem stärkeren negativen Effekt führten.

Allerdings bestand in OSCE II ein inverser Unterschied, dort waren die Proband*innen mit Deutsch als Muttersprache in der regulären Station signifikant besser als die Proband*innen mit einer anderen Muttersprache und in der Studienstation bestand kein Unterschied.

Es lässt sich also keine eindeutige Aussage zu den Effekten der Muttersprache auf die Leistungen treffen. Möglicherweise war die Variable Muttersprache nicht aussagekräftig genug. Die Gruppe von Proband*innen mit einer anderen Muttersprache als Deutsch kann sehr inhomogen sein und die einzelnen Proband*innen unterscheiden sich hinsichtlich ihrer Deutschkenntnisse möglicherweise so stark, dass ein Vergleich nicht möglich bzw. sinnvoll ist. Außerdem war der Anteil von Proband*innen mit einer anderen Muttersprache sowohl bei OSCE I, als auch OSCE II sehr gering. Womöglich war die Fallzahl zu gering um einen Effekt berechnen zu können.

5.2.8. Effekt der Telefonanrufe abhängig von den Variablen Sport, Geschwister, Videospiele

Wie aus Studien bekannt gibt es viele verschiedene Faktoren, die Einfluss auf den Effekt von Unterbrechungen und Ablenkungen haben. Einige sind individuell unterschiedlich. Dazu zählen unter anderem, ob die Proband*innen bisher mit Unterbrechungen Vorerfahrungen gemacht haben und ob sie sich Strategien im Umgang mit Unterbrechungen angeeignet haben.^{63 97}

Daher wurden einige individuelle Merkmale und Verhaltensweisen der Proband*innen in den Fragebögen abgefragt, die mit einem regelmäßigen Umgang mit Multitasking, Unterbrechungen und Ablenkungen assoziiert sein könnten oder bei welchen ein positiver Effekt auf kognitive Leistungen und Multitasking nachgewiesen werden konnte^{141–146}.

Dazu zählte die Anzahl von Geschwistern, wie viele Stunden die Proband*innen Sport pro Woche machen und ob sie Videospiele spielen.

So gibt es Studien, die darauf hinweisen, dass Sportler*innen in kognitiven Testungen besser abschneiden als eine Kontrollgruppe, die weniger Sport betreibt^{141–143}. Die in den Studien getesteten kognitiven Leistungen bezogen sich unter anderem auf die Aufmerksamkeit und das Arbeitsgedächtnis, welche beide für den Umgang mit Unterbrechungen relevant sind^{126,141–143,147}.

Auch konnten experimentelle Studien zeigen, dass ein Training mit Videospielen die kognitiven Funktionen der Proband*innen, unter anderem die Fähigkeit von Multitasking, verbessern konnte^{144–146}.

Auch wenn es diesbezüglich noch keine eindeutigen Hinweise in der Literatur gibt, wurde zudem die Anzahl der Geschwister erhoben, in der Annahme, dass dies ebenfalls mit Vorerfahrungen mit Unterbrechungen assoziiert sein könnte.

Die Korrelationsanalysen zeigten, dass es keinen Unterschied in der erreichten Punktezahl in der Studienstation und regulären Station zwischen den Proband*innen mit den entsprechenden Merkmalen oder ohne die Merkmale gab.

Der Umgang mit Unterbrechungen und Ablenkungen und deren Auswirkungen sind ein komplexes Forschungsfeld, in denen es schwer ist einzelne Faktoren herauszuarbeiten, die einen Einfluss haben könnten. Möglicherweise wurden die Merkmale über den Fragebogen zu ungenau erfasst oder der Anteil der Proband*innen war insgesamt zu gering um einen Effekt zu detektieren.

5.3 Konsequenzen von Unterbrechungen durch Telefonanrufe für die Patient*innensicherheit und ärztliche Tätigkeit

Die Hypothese dieser Arbeit, dass Unterbrechungen durch Telefonanrufe zu schlechteren Leistungen in einer OSCE-Prüfung führen, wurde bestätigt. Erstmals konnte mit dieser Studie ein negativer Effekt durch Unterbrechungen durch Telefonanrufe auf ärztliche Tätigkeiten außerhalb des Operationssaals mit einem experimentellen Ansatz nachgewiesen werden.

Alltägliche ärztliche Aufgaben wie die körperliche Untersuchung eines Patienten/ einer Patientin, das Durchführen eines Anamnesegesprächs oder kleinere Eingriffe, wie Blutabnahmen oder das Legen einer Venenverweilkanüle fallen in beinahe allen Fachrichtungen an und nehmen bis zu 46% der Arbeitszeit von Ärzt*innen in Anspruch ^{31,148}. Sie werden sowohl von Assistenzärzt*innen durchgeführt als auch von Medizinstudierenden im letzten Jahr ihrer Ausbildung (Praktisches Jahr). Mit dem Untersuchen der Effekte von Unterbrechungen durch Telefonanrufe auf alltägliche ärztliche Tätigkeiten konnte mit dieser Studie eine wichtige Forschungslücke geschlossen werden.

In der Auswertung der einzelnen Items der Checkliste wurde deutlich, dass sowohl in OSCE I, als auch in OSCE II sicherheitsrelevante Schritte vergessen wurden. In OSCE I vergaßen die Proband*innen Desinfektionsmittel, Tupfer, Abwurf und Stauschlauch zu richten. Außerdem bekamen sie signifikant weniger Punkte für das Handschuhe anziehen und die vorherige Händedesinfektion. Die Händedesinfektion ist ein wichtiger Schritt in der Vermeidung von Infektionen und der Übertragung von Erregern. Zudem sind Handschuhe eine wichtige Maßnahme, um medizinisches Personal vor der Übertragung von möglichen Infektionen zu schützen.

In OSCE II erreichten die Proband*innen für die körperliche Untersuchung und Anamnese in der Studienstation signifikant weniger Punkte. Wenn die Anamnese und körperliche Untersuchung eines Patienten/ einer Patientin nicht vollständig sind, können unter Umständen wichtige Informationen und Befunde fehlen. Dies kann z.B. zu medizinischen Fehleinschätzungen führen. In der Studie gab es jedoch erfreulicherweise keinen signifikanten Unterschied in der Anzahl richtiger Diagnosen. Allerdings erhielten die Proband*innen für die Anamnese Punkte Vorerkrankungen und -operationen sowie Allergien und Medikamente in der Studienstation signifikant weniger Punkte. Diese Informationen können sicherheitsrelevant für den Patienten/ die Patientin sein, beispielsweise wenn Allergien gegen Medikament oder Kontraindikationen für Therapien bestehen, eine medikamentöse Antikoagulation eingenommen wird oder der Patient voroperiert wurde. Falls der Patient/ die Patientin operiert werden soll, sind diese Informationen beispielsweise von hoher Relevanz.

Es konnte demnach nicht nur gezeigt werden, dass die Proband*innen durch die Telefonanrufe in der Studienstation schlechter abschnitten, sondern auch, dass dies potenziell negative Konsequenzen für die Patient*innensicherheit haben kann.

Damit bestätigt die Studie bereits bestehende Erkenntnisse, dass Unterbrechungen und Ablenkungen von medizinischem Personal nicht nur zu einer hohen Arbeitsbelastung^{5,14–16,23}, zu mehr Frustration^{13,23}, einem hohen Stresslevel^{9–16} und mehr Überarbeitung und Erschöpfung²⁴ sondern auch zu einer schlechteren Behandlungsqualität⁹, zu mehr Fehlern, beispielsweise in der Medikation^{19,20} oder während Operationen^{21,22} und zu Informations- und Fokusverlust führen können¹³.

5.4 Lösungsansatz: Sensibilisierung von medizinischem Personal und Training im Umgang mit Unterbrechungen und Ablenkungen

Es bestehen verschiedene Optionen, welche Konsequenzen aus dieser Studie gezogen werden können. Telefonanrufe sind eine sehr häufige Ursache von Unterbrechungen und ein wichtiges Kommunikationsmittel in Kliniken. Es muss demnach am ehesten ein angemessener Umgang mit den Unterbrechungen gefunden werden, um potenziell negative Effekte zu vermeiden.

Ein möglicher Ansatz, ist eine Sensibilisierung des medizinischen Personals für mögliche negative Konsequenzen zu schaffen und ein Etablieren von Trainings im Umgang mit Unterbrechungen.

Eine Sensibilisierung scheint besonders wichtig, da Studien zeigten, dass Ablenkungen durch Telefone nicht nur zu mehr Fehlern, sondern auch zu einer falschen Einschätzung der eigenen Fähigkeiten führen können^{70,82}. Dies zeigte sich sowohl in einer Studie zum Autofahren als auch in der Studie von Persoon et al. während einer simulierten Operation^{70,82}. Nur zehn Prozent der Proband*innen gaben an, dass sie sich durch die Ablenkungen belastet oder eingeschränkt fühlten, obwohl sie nachweislich mehr Fehler machten⁸². Demnach können Ablenkungen nicht nur zu mehr Fehlern führen, sondern auch zu einer falschen Selbstwahrnehmung. Dies ist insbesondere fatal, da das Risiko für Fehler durch Ablenkungen steigt und im Gegenteil eine erhöhte Aufmerksamkeit bestehen sollte, um das Risiko zu reduzieren.

Tatsächlich wird das Praktizieren und Üben von Achtsamkeit und Aufmerksamkeit von medizinischem Personal durch manche Autor*innen gefordert, um beispielsweise Medikationsfehler zu reduzieren¹⁴⁹. Teilweise scheint dies medizinischem Personal bereits zu gelingen, so gaben in der Studie von Augenstein et al. die befragten Ärzt*innen in der

Notaufnahme bei einer hohen Anzahl an Multitasking ein gesteigertes Situationsbewusstsein und eine gesteigerte Situationsaufmerksamkeit an (im englischen als situation awareness bezeichnet) ¹¹. Weigl et al. stellten hingegen fest, dass eine hohe Anzahl von Unterbrechungen mit einem geringeren Situationsbewusstsein des medizinischen Personals in Notaufnahmen assoziiert sei, dies aber je nach Quelle der Unterbrechung unterschiedlich ausgeprägt sei ⁴¹. Unter anderem würden Unterbrechungen durch Telefone und Pager zu einem geringeren Situationsbewusstsein führen ⁴¹. Eine weitergehende Aufklärung und Training von sogenannten nicht-technischen Kompetenzen erscheint sinnvoll. Es gibt bereits verschiedene Studien, in denen unterschiedliche Modelle für die Lehre eingesetzt und untersucht wurden. In diesen wurde u.a. Kommunikationsfähigkeit, Teamarbeit, Achtsamkeit und Strategien im Umgang mit Unterbrechungen und Multitasking gelehrt und beurteilt.

In einer Studie von Hochberg et al. durchliefen die Proband*innen (Chirurgie-Assistenzärzt*innen) beispielsweise vor und nach einem Kompetenz-Training mit dem Fokus auf Professionalität und Kommunikation eine OSCE-Prüfung. Sie waren nach dem Training signifikant besser. Es wurde unter anderem auch das Management von Telefonanrufen gelehrt und geprüft ¹¹⁹.

In einer anderen Studie erhielten Assistenzärzt*innen eine online Lerneinheit zu der Durchführung eines Übergabegesprächs und zu dem Umgang mit Ablenkungen. Anschließend durchliefen sie eine Übergabe-Simulation unter verschiedenen Bedingungen. Im Anschluss gaben 80% der Proband*innen an, effektivere Übergabegespräche führen zu können ¹⁵⁰. Eine ähnliche Studie setzte den Fokus auf die Vorbereitung von Radiologie Assistenzärzt*innen auf ihre Dienste. Sie erhielten unter anderem Trainingseinheiten zu dem Umgang mit Unterbrechungen und Telefonanrufen. Dadurch fühlten sich 89% der Proband*innen besser auf ihre Dienste vorbereitet ¹⁵¹.

Auch für Assistenzärzt*innen der Urologie wurden simulierte Visiten mit Schauspieler*innen und anschließendem Feedbackgesprächen als Lehrmethode erfolgreich eingesetzt. In diesen wurden auch Ablenkungen durch Telefonanrufe simuliert ^{152,153}.

Eine ähnliche Studie mit Medizinstudierenden zeigte anhand einer Kontrollgruppe, dass Feedbackgespräche nach einer simulierten Visite die Anzahl von Fehlern in einer später durchlaufenen simulierten Visite reduzierten. Ein Lerneffekt konnte in beiden Gruppen nachgewiesen werden, war aber in der Interventionsgruppe mit Feedbackgesprächen ausgeprägter ¹⁵⁴. Auch in der Lehre von Krankenpflegeschüler*innen konnte anhand eines Rollenspiels eine Sensibilisierung und ein besseres Verständnis für mögliche Konsequenzen von Unterbrechungen und Ablenkungen während der Vorbereitung von Medikamenten geschaffen werden ¹³⁸.

Zusammenfassend lässt sich feststellen, dass es bereits vielversprechende Ansätze in der medizinischen Lehre gibt und bisherige Studien positive Effekte nachweisen konnten.

Es wäre denkbar an der medizinischen Fakultät der Universität zu Köln zur Verbesserung der nicht-technischen Kompetenzen dauerhaft in den OSCE-Prüfungen eine Station zu etablieren, in welcher die Studierenden mit Ablenkungen und Unterbrechungen durch Telefonanrufe konfrontiert werden. Ein anschließendes Seminar zur Nachbesprechung sollte dabei Teil der Lerneinheit sein. Die Ergebnisse von OSCE I und OSCE II könnten zum Evaluieren genutzt werden.

Bei medizinpädagogischen Ansätzen, wie einem entsprechenden Seminar, sollte eine Berücksichtigung der in den Abschnitten 5.2 herausgearbeiteten Einflussfaktoren auf den Effekt von Unterbrechungen erfolgen.

So ist insbesondere der Zeitpunkt einer Unterbrechung relevant. In der Mitte und am Ende einer Tätigkeit besteht ein erhöhtes Risiko für negative Effekte durch Unterbrechungen.

Ebenso sollte die Komplexität der primären und sekundären Aufgabe berücksichtigt werden. Handelt es sich beispielsweise um eine sehr komplexe Tätigkeit, bei welcher ein Arzt/ eine Ärztin unterbrochen wird, sollte erwogen werden, das Telefonat auf einen späteren Zeitpunkt zu verschieben oder so kurz wie möglich zu telefonieren. Ebenso wenn es sich um ein komplexes Gespräch/Telefonat handelt.

Telefonanrufe können sowohl zu einer Dual-Task führen (während des Telefongesprächs setzt der Arzt oder die Ärztin beispielsweise die Dokumentation am Computer fort oder die Verordnung eines Medikaments) oder als Task-Switch (der Arzt oder die Ärztin unterbricht während des Telefonats die primäre Tätigkeit, z.B. ein Anamnesegespräch und führt es anschließend fort). Je nach Komplexität der Aufgaben, Zeitpunkt und Vorerfahrung können so verschiedene Strategien und Entscheidungen getroffen werden.

Es ist davon auszugehen, dass viele Ärzt*innen bereits Strategien im Umgang mit Unterbrechungen und Ablenkungen entwickelt haben. Ein Fokus in der Weiterbildung sollte daher insbesondere auf jungen Assistenzärzt*innen und Medizinstudierenden liegen. Wie im Abschnitt 5.2.5 ausgeführt, sind Vorerfahrungen ein relevanter Faktor im Umgang mit Unterbrechungen und Ablenkungen. Besonders Assistenzärzt*innen mit wenig Berufserfahrung scheinen gefährdet, durch Ablenkungen und Unterbrechungen mehr Fehler zu machen.

Besonders anschaulich wird die Umsetzung der verschiedenen Maßnahmen bei den Bemühungen, Fehler beim Richten von Medikamenten durch Pfleger*innen durch

Ablenkungen und Unterbrechungen zu reduzieren. So wurden, wie oben bereits beschrieben, Rollenspiele in der Ausbildung von Pflegeschüler*innen eingesetzt ¹³⁸. In einer anderen Studie wurden verschiedene Maßnahmen ergriffen, um die Anzahl an Unterbrechungen und Ablenkungen zu reduzieren. Zu den Maßnahmen zählen unter anderem das Etablieren einer Checkliste und das Markieren einer bestimmten Zone, in welcher die Medikamente gerichtet werden und Störungen vermieden werden sollten ¹³⁷. Durch diese Maßnahmen konnte die Anzahl von Unterbrechungen und Ablenkungen, sowie Medikationsfehler signifikant reduziert werden ¹³⁷.

Das Etablieren von Hilfsmitteln, wie beispielsweise einer Checklisten, könnte auch bei einer Anamneseerhebung hilfreich sein. Möglicherweise hätten die Proband*innen in OSCE II weniger vergessen, wenn ihnen eine entsprechende Checkliste zur Verfügung gestanden hätte. In einigen Kliniken werden bereits standardisierte Anamnesebögen eingesetzt, in anderen Kliniken gibt es kein einheitliches Vorgehen.

5.5 Lösungsansatz: Human Factor und systemische Veränderungen

Zunehmende Beachtung in der wissenschaftlichen Auseinandersetzung mit Patient*innensicherheit finden zudem Konzepte aus der Luft- und Raumfahrtindustrie. Dort wurde sehr früh erkannt, dass der Mensch als potenzielle Fehlerquelle („Human Factor“) in Sicherheitskonzepten berücksichtigt werden sollte. Es wird gefordert, dass Human factor Forschung in der Medizin einen ähnlich hohen Stellenwert, wie in der Luftfahrtindustrie erhalten sollte ^{155–157}. Teilweise wurden bereits Ansätze auf Gesundheitssysteme übertragen, wie Kommunikation- und Teamarbeit Trainings ¹⁵⁸. Auch die im vorherigen Abschnitt beschriebenen Maßnahmen, um Fehler bei dem Richten von Medikamenten durch Pfleger*innen zu reduzieren, basierten auf Konzepten der Luftfahrtindustrie ¹³⁷. Ferner konnten Studien zeigen, dass das Konzept „steriles Cockpit“ auch bei der Durchführung von endoskopischen Untersuchungen und Eingriffen die Anzahl von Unterbrechungen und Ablenkungen reduzieren konnte ^{159,160}.

Human Factor Forschung setzt jedoch nicht nur das Individuum in das Zentrum seiner Forschung und Optimierungsbemühungen, sondern konzentriert sich vor allem auf das System, in welchem die Individuen arbeiten ¹⁵⁵. Im Sinne der Human Factor Forschung sollte das gesamte System betrachtet und optimiert werden ¹⁵⁵.

Ein möglicher systemischer Lösungsansatz in Bezug auf Unterbrechungen und Ablenkungen durch Telefonanrufe ist der Versuch unnötige Unterbrechungen zu reduzieren. Studien zeigten, dass Pager und Telefonanrufe selten aufgrund von Notfällen getätigt wurden und ein nicht unerheblicher Anteil der Anrufe und Pager von Ärzt*innen als unnötig oder vermeidbar eingeschätzt wurden ^{13,161–164}. Es wäre wünschenswert, diese Unterbrechungen zu vermeiden. Vereinzelt wurden bereits Maßnahmen mit diesem Ziel untersucht. In einer

radiologischen Abteilung wurde für drei Monate ein Telefon Triage System etabliert. Telefonanrufe wurden zunächst an einen Mitarbeiter/ eine Mitarbeiterin geleitet, der entschied, ob die Telefonanrufe an den diensthabenden Arzt/ die diensthabende Ärztin weitergeleitet werden sollen oder nicht. Dies führte zu einer Reduktion der Unterbrechungen durch Telefonanrufe um 71% ¹⁶⁵. In anderen Studien wurde in Notaufnahmen ein schriftliches elektronisches Kommunikationsmittel etabliert, um die Anzahl von Pagern und Telefonanrufen zu reduzieren. Dies führte zu einer Reduktion der Anrufe durch Pfleger*innen, die nicht dringlich waren, und zu weniger Arbeitsflussunterbrechungen der Ärzt*innen ¹⁶⁶. In Ernst et al. Studie wurde die Einführung eines Freisprech-Kommunikation Geräts, über welches der Nutzer/ die Nutzerin per Sprachbefehl selbst Anrufe anfordern und eingehende Anrufe annehmen oder ablehnen kann, evaluiert. Durch die neuen Geräte, ergänzend zu zuvor genutzten Telefonen und Pagern, waren die kommunikationsbedingten Unterbrechungen signifikant kürzer, allerdings gab es mehr Unterbrechungen pro Stunde ¹⁶⁷. Da auch die Dauer einer Unterbrechung einen Einfluss auf die potenziell negativen Effekte haben kann, ist dies dennoch als positiv zu bewerten ⁶³. Interessant wäre in dem Fall also gewesen, ob es durch die erhöhte Frequenz aber geringerer Dauer von Unterbrechungen zu einem positiven oder negativen Effekt auf die Konzentrationsfähigkeit und Patient*innensicherheit kam.

In der Bewertung der Studienergebnisse sollte jedoch berücksichtigt werden, dass Unterbrechungen und Ablenkungen von medizinischem Personal nicht ausschließlich negative Folgen haben ^{73,74}. Häufig sind diese unerlässlich, um eine gute Patientenversorgung gewährleisten zu können und auch in akuten Notfallsituationen schnell reagieren zu können. Beispielsweise gaben in der Studie von Armendariz et al. 29% der Befragten an, dass die Unterbrechungen während einer Visite auch positive Effekte haben könnten, v.a. wenn es um eine Akutsituation in der Patientenversorgung gehe ¹³.

Entscheidend ist also, zwischen den relevanten Telefonanrufen zu unterscheiden und dabei abzuwägen, inwiefern durch die Unterbrechung die Patientenversorgung verbessert wird bzw. sich diese durch die Folgen der Ablenkung, mit einem erhöhten Risiko für Fehler, verschlechtert.

In einem systemischen Ansatz sollten die Telefonanrufe reduziert werden, die nicht unbedingt nötig sind. Zum anderen müssen die Situationen oder Tätigkeiten identifiziert werden, die besonders anfällig für Fehler durch Unterbrechungen sind. Auf letztere kann dann durch Aufklärung und Trainings besser vorbereitet werden.

5.6 Diskussion der Methodik und Limitationen

In dieser randomisierten Studie wurden 161 Proband*innen in OSCE I und 147 Proband*innen in OSCE II eingeschlossen.

Damit erreichte die Studie ein Alleinstellungsmerkmal im Vergleich zu ähnlichen experimentellen Studien zu den Effekten von Ablenkungen und Unterbrechungen auf ärztliche Tätigkeiten, die deutlich weniger Proband*innen einschließen konnten. In den meisten Studien waren es kleinere Proband*innenzahlen von 12 - 41 Personen ^{25–28,80,81,83–95,121,168}.

Nur Persoon et al. und Monteiro et al. Studien hatten eine größere Proband*innenzahl ^{82,96}. Persoon et al. konnten in ihrer Studie mit 86 Proband*innen mehr Fehler durch Unterbrechungen während einer virtuellen endoskopischen Operation nachweisen ⁸². Monteiro et al. hingegen konnten in ihrer Studie mit 198 Proband*innen keine höhere Anzahl von Diagnosefehlern, aber eine längere Bearbeitungsdauer durch Unterbrechungen nachweisen ⁹⁶.

Außerdem ist die vorliegende Arbeit nach dem aktuellen Stand der Literaturrecherche die einzige Studie, die in dieser Form den Effekt von Telefonanrufen auf die Durchführung alltäglicher ärztlicher Tätigkeiten fachbereichs-übergreifend prüfte.

Limitationen dieser Arbeit sind vor allem durch verschiedene Aspekte des Studiendesigns begründet und sind im Folgenden aufgeführt.

5.6.1. Studiendesign und interne Validität

Für die vorliegende Studie wurde ein time series, within-group Design gewählt, es wurden die Leistungen derselben Proband*innen unter verschiedenen Bedingungen untersucht ¹²⁰. Es gab keine Kontrollgruppe, da die Proband*innen selbst als Kontrolle fungierten. Ein Vorteil des within-subject Design ist, dass Unterschiede in den Leistungen der Proband*innen unter unterschiedlichen Bedingungen nicht auf individuelle Unterschiede zwischen den Proband*innen einer Interventionsgruppe und Kontrollgruppe zurückgeführt werden können ¹²⁰. Teilnehmerbezogene Bedrohungen für die interne Validität ergeben sich bei dem vorliegenden Studiendesign daher nicht, insbesondere keine Unterschiede in der Selektion der Proband*innen. Daher wird die statistische Power meist höher als in between-subject designs bewertet ¹⁶⁹.

Dennoch ergeben sich aus dem Studiendesign andere Risiken für die interne Validität, welche in den folgenden Abschnitten diskutiert werden.

5.6.2. Bedrohungen für die interne Validität: Übertragungseffekte, Übungseffekte, Sensibilisierung

Bedrohungen für die interne Validität, die sich aus dem gewählten Studiendesign in der vorliegenden Studie ergeben, sind Übertragungseffekte (carry-over Effect) und Übungseffekte (Practice Effect) ^{169,170}. Übertragungseffekte bestehen dann, wenn eine Maßnahme oder Intervention einen länger anhaltenden Einfluss auf die Proband*innen hat und auch in der folgenden Messung das Ergebnis beeinflusst ¹⁷⁰.

Übungseffekte sind relevant, wenn die erhobenen Daten von bestimmten Fertigkeiten der Proband*innen abhängen, die durch Übung verbessert werden können ^{169,170}. Im Falle der Studie war dies relevant, da die Proband*innen zwei mal die gleiche Aufgabe bearbeiten mussten. Ein Lerneffekt innerhalb der Studie und damit eine Verfälschung der Ergebnisse war demnach nicht auszuschließen.

Um dies zu verhindern hat ein Teil der der Proband*innen erst die Studienstation und dann die reguläre Station durchlaufen und dann andersherum. Die Zuteilung erfolgte zufällig. So sollte ausgeschlossen werden, dass ein Lerneffekt sowohl von der Studienstation auf die reguläre Station bestand, als auch andersherum. In OSCE II wurden zudem die Szenarien gewechselt, sodass die Proband*innen unterschiedliche Krankheitsbilder geprüft wurden. Auch ein Übertragungseffekt von einer Station auf die andere sollte so ausgeglichen werden.

Eine weitere Bedrohung für die interne Validität lag bei der Sensibilisierung der Proband*inne ^{169,170}. Es ist möglich, dass die Proband*innen im Verlauf der Studie eine Hypothese entwickelten, was die Fragestellung der Studie sein könnte und dies ihr Verhalten beeinflusst haben könnte ^{169,170}.

5.6.3. Bedrohung für die interne Validität: Instrumentation

Eine weitere Bedrohung der internen Validität nach Casell et al. ergibt sich aus „Instrumentation“ ^{120,171}. Diese Bedrohung besteht, wenn die Messmethoden zur Datenerhebung während des Studienablaufs wechseln oder die Prüfer*innen mit der Bewertung vertraut werden und ihr Bewertungssystem anpassen ^{120,171}.

Um im Verlauf der Studie identische Messmethoden anzuwenden, wurde durchgehend die gleiche Checklisten verwendet. Um zu verhindern, dass sich die Prüfer*innen in den Bewertungen unterscheiden, erfolgte zuvor eine gemeinsame Schulung.

Um dennoch auszuschließen, dass es signifikante Unterschiede in den Bewertungen der verschiedenen Prüfer gab, wurde eine univariate Varianzanalyse (ANOVA) der Prüfer*innen durchgeführt. Außerdem ein Korrelationstest mit den Ergebnissen der Prüfer*innen der Universität. Es konnte kein Unterschied in der Bewertung zwischen den verschiedenen

Prüfer*innen der Studie nachgewiesen werden. Die Bewertungen aus der Studie korrelierten zudem mit den Bewertungen der Prüfer*innen der Universität.

Weitere Prozedere bezogene Risiken für die interne Validität ergaben sich durch die Schauspielpatient*innen. In der regulären Station wurden professionelle Schauspieler*innen eingesetzt, in der Studienstation hingegen Doktorand*innen der Arbeitsgruppe. Möglicherweise fiel es den Proband*innen dadurch leichter oder schwerer in der Studienstation alle Punkte zu erreichen. Beides hätte Einfluss auf das Gesamtstudienergebnis. Um dem entgegen zu wirken erhielten die Doktorand*innen gemeinsam mit den professionellen Schauspieler*innen eine Schulung und ein Skript.

Da die Leistungen aus der Studienstation nicht für die Ergebnisse der offiziellen OSCE-Prüfung, in welche die Studie eingebettet war, relevant waren ist zudem denkbar, dass sich die Studierenden weniger Mühe in der Bearbeitung der Aufgabe gaben und es dadurch zu einer schlechteren Leistung der Proband*innen kam. Wäre die Studie unabhängig von den offiziellen OSCE-Prüfungen durchgeführt worden, hätte dies vermieden werden können. Allerdings konnte so die Logistik der OSCE-Prüfungen genutzt werden und eine sehr große Anzahl von Proband*innen in die Studie eingeschlossen werden.

In OSCE II unterschieden sich zudem die reguläre Station und die Studienstation in den Räumlichkeiten. Die reguläre Station wurde in einem kleinen Raum durchgeführt, in dem sich nur der*die Schauspieler*in befand. Der*die Prüfer*in befand sich in einem Nebenraum und konnte über eine einseitig verspiegelte Glasscheibe und Kopfhörer die Prüfung verfolgen. In der Studienstation befand sich der*die Prüfer*in jedoch im gleichen Raum. Dies könnte zu einer zusätzlichen Ablenkung der Proband*innen geführt haben, sodass der negative Effekt auf die Leistungen nicht nur auf die Ablenkungen durch die Telefonanrufe zurückgeführt werden könnte. Allerdings befanden sich auch in anderen Stationen der OSCE-Prüfung die Prüfer*innen im selben Raum, sodass die Situation für die Proband*innen nicht gänzlich neu war.

Alternativ hätte ein between-group design gewählt werden können, in dem einige Proband*innen die Station ohne Telefonanrufe durchlaufen und andere nur die Station mit Telefonanrufen durchlaufen hätten. Für ein between-group Designs ist jedoch eine größere Anzahl von Proband*innen nötig, es bestehen andere Risiken für die interne Validität, wie durch die Selektion, und die Umsetzung ist meist wesentlich aufwendiger ^{120,169–171}.

Vermutlich aus diesen Gründen wurde auch in den meisten bisherigen experimentellen Studien, die den Effekt von Unterbrechungen und Ablenkungen auf Tätigkeiten von

medizinischem Personal untersuchten, ein within-group Design genutzt. Persoon et al. setzte hingegen eine Kontrollgruppe ein und führte einen between-group Vergleich durch⁸².

Durch das gewählte Studiendesign konnten zudem optimal die lokalen Gegebenheiten der OSCE-Prüfungen an der Universität zu Köln genutzt und Proband*innen akquiriert werden.

5.6.4. Kritische Reflektion der Methodik in Bezug auf die Szenarien in OSCE II

Insgesamt wurden in OSCE II drei unterschiedliche Szenarien (Divertikulitis, Appendizitis, Cholezystitis) in verschiedenen Kombinationen eingesetzt. In der Ergebnisanalyse wurde deutlich, dass der Effekt der Unterbrechungen (gemessen anhand der unterschiedlichen Punktwerte in der Station mit Telefonanrufen und ohne Telefonanrufe) je nach Szenarien Kombination unterschiedlich war.

Der Unterschied in den erreichten Punktezahlen zwischen der Studienstation und der regulären Station war je nach Szenarienkombination unterschiedlich stark ausgeprägt.

Im Vergleich der durchschnittlich erreichten Punktezahlen in den verschiedenen Szenarien, unabhängig von einer Unterbrechung, unterschieden sich die Szenarien jedoch nicht signifikant. Es konnte lediglich der Trend aufgezeigt werden, dass die Proband*innen im Appendizitis Szenario besser abschnitten. Da in der mixed ANOVA jedoch ein unterschiedlicher Effekt der Unterbrechungen je nach Szenario festgestellt wurde, ist es möglich, dass sich die Szenarien im Anforderungsprofil, zumindest wenn gleichzeitig Unterbrechungen durch Telefonanrufe auftraten, unterschieden.

Es stellt sich die Frage, inwiefern dadurch ein Einfluss auf das Gesamtstudienereignis besteht. Da es jedoch sowohl Kombinationen gab, in denen die Punktezahldifferenz höher war, als auch solche, in denen diese niedriger war, kann der Einfluss auf das Gesamtergebnis möglicherweise wieder ausgeglichen werden.

Außerdem stellt sich grundsätzlich die Frage, ob eine Vergleichbarkeit der Leistungen eines Probanden/ einer Probandin in der Studienstation und in der regulären Station gegeben ist, wenn zwei unterschiedliche Szenarien geprüft wurden, wie es in OSCE II der Fall war.

Betrachtet man die Methodik bisheriger Studien zu den Effekten von Unterbrechungen und Ablenkungen lassen sich unterschiedliche Herangehensweisen feststellen.

In Studien, in welchen der Effekt von Unterbrechungen und Ablenkungen auf die Leistungen im Rahmen einer simulierten Operation getestet werden sollten, handelte es sich meistens ebenfalls um ein time series, within-group Design. Die Proband*innen führten die gleiche Operationsaufgabe unter unterschiedlichen Bedingungen durch, wie es auch in OSCE I in der vorliegenden Studie der Fall war. Entweder sie durchliefen zu Beginn der Studie Übungen, bis

ein bestimmtes Niveau erreicht wurde, um dann einen Ausgangswert unter ungestörten Bedingungen zu erheben und im Anschluss eine Simulation mit Störungen zu durchlaufen^{27,28,81,86,89,92} oder sie sollten meist nach einigen Übungsdurchläufen in einer randomisierten Reihenfolge die Operation mit oder ohne Störungen durchführen^{25,26,80,83–85,87,90,94}.

In der vorliegenden Studie wurde die letztere Variante gewählt, in der die Proband*innen in zufälliger Reihenfolge die Stationen durchliefen, um einen Lerneffekt vorzubeugen.

In Studien, die die Beurteilung der Proband*innen von Patientenfällen unter unterschiedlichen Bedingungen evaluierten, wurden die Fälle wie in der vorliegenden Studie in OSCE II, variiert. In der Studie von Monteiro et al. sollten die Proband*innen insgesamt zwanzig unterschiedliche schriftliche Patientenfälle bearbeiten. In der Hälfte der Fälle wurden sie dabei unterbrochen⁹⁶. Es wurden also nicht die exakt gleichen Patientenfälle für den Vergleich der Leistungen mit oder ohne Unterbrechungen genutzt, wie auch in dieser Studie nicht genau die gleichen Szenarien miteinander verglichen wurden. In Woods et al. Studie sollten die Proband*innen einen venösen Zugang legen, einer Patientenanamnese zuhören und Fragen dazu beantworten und im Anschluss beide Aufgaben gleichzeitig erledigen. In der Studie wurde ähnlich wie in der vorliegenden Studie die Patientengeschichte variiert, um einen Lerneffekt zu vermeiden¹²¹. Dies erscheint sinnvoll, da beispielsweise in der Studie von Evans et al. ein Lerneffekt in den korrekten Antworten nachgewiesen werden konnte, als die Proband*innen wiederholt durch die gleichen Patientenfall-Fragen abgelenkt wurden⁸⁸.

Zusammenfassend lässt sich feststellen, dass eine Variation der Patientenfälle ein adäquates Mittel sein kann, um einen Lerneffekt zu vermeiden. Eine mögliche Abweichung der Leistungen und Ergebnisse durch beispielsweise unterschiedliche Anforderungen der Fälle muss dabei jedoch in Kauf genommen werden.

Da in dieser Studie eine Zuteilung zu den Szenarienkombinationen zufällig erfolgte und insgesamt sechs unterschiedliche Kombinationen geprüft wurden scheint ein relevanter Einfluss auf das Gesamtstudienresultat unwahrscheinlich. Dennoch wäre angesichts der unterschiedlichen Anforderungsprofile der Szenarien, welche sich in der Ergebnisanalyse zeigte, wünschenswert gewesen alle Szenarien Kombinationen in der gleichen Anzahl zu prüfen, um potenzielle Einflüsse besser ausschließen zu können.

Letztlich konnte in der vorliegenden Studie sowohl in OSCE I, bei der Prüfung der gleichen Aufgabe, als auch in OSCE II bei der Prüfung unterschiedlicher Aufgaben ein Effekt der Ablenkungen nachgewiesen werden.

5.6.5. Externe Validität

Die externe Validität befasst sich nach Creswell et al. mit der Frage, ob Ergebnisse einer experimentellen Studie verallgemeinert und auf andere Personen, Situationen und Zeitpunkte übertragen werden können ^{120,171}

Dies wird in den folgenden beiden Abschnitten diskutiert.

5.6.6. Übertragbarkeit der Ergebnisse der OSCE-Prüfung auf den ärztlichen Alltag

Zur Überprüfung, ob Unterbrechungen durch Telefonanrufe auf medizinisches Personal einen Einfluss haben, wurde in dieser experimentellen Studie das Prüfungsformat OSCE gewählt. In einer OSCE-Prüfung werden fächerübergreifende realistische Situationen aus dem Arztalltag simuliert.

Gleichwohl ist zu diskutieren, ob die Erkenntnisse aus der OSCE-Prüfung auf den klinischen Alltag von Ärzt*innen übertragbar sind. Wie nah sind OSCE-Prüfungen tatsächlich an der Realität und können diese simulieren? Wie zuverlässig können die Leistungen von Studierenden geprüft werden?

Die Gütekriterien Validität und Reliabilität wurden für das Prüfungsformat OSCE vielfach evaluiert. Viele der Studien kamen zu zufriedenstellenden Ergebnissen ^{100,107,110,112,113,116}. Hodges et al. identifizierte verschiedene Faktoren, die auf die Validität und Reliabilität einer OSCE-Prüfung Einfluss haben können ¹⁷². Dazu zählen unter anderem der soziokulturelle Kontext, in dem die Prüfung stattfindet, und die Eigenschaften und Voraussetzungen der geprüften Gruppe ¹⁷². Auch die Leistung der Schauspielpatient*innen, die Unabhängigkeit der Prüfer*innen, die verwendeten Bewertungsbögen, das Design der Stationen und die Anzahl der Stationen haben einen Einfluss ^{99,105,172–176}. Bei einer sorgfältig geplanten und durchgeführten OSCE-Prüfung scheint eine hohe Reliabilität und Validität dieses Prüfungsformats möglich ^{100,107,110,112,113,116}. In der Vergangenheit haben Prüfungsteilnehmer*innen verschiedener OSCE-Prüfungen in der Medizin die OSCE als authentische Prüfungsform bewertet, in welcher auf eine angemessene Art und Weise klinisch praktische Fertigkeiten geprüft werden ^{106,113,174,177}. Eine Befragung der Proband*innen im Anschluss an die Prüfung wurde in dieser Studie nicht durchgeführt. Allerdings wird die OSCE-Prüfung mit den entsprechenden Stationen bereits seit vielen Jahren an der medizinischen Fakultät Köln als Prüfung eingesetzt und mit einer weitreichenden Erfahrung der Organisator*innen sowie der Prüfer*innen umgesetzt. Weiterhin kann von einer Augenscheinvalidität („face validity“) der OSCE Studienstationen ausgegangen werden, da allgemein bekannt ist, dass das Legen eines venösen Zugangs und die Durchführung eines

Anamnesegesprächs und einer klinischen Untersuchung auch in der Realität ärztliche Aufgaben sind.

Als Kriterium der Validität kann außerdem hinzugezogen werden, dass die Proband*innen mit einem klinischen Erfahrungsvorsprung (einer Berufsausbildung im medizinischen Bereich) besser in den OSCE I Prüfungen waren, als die die Proband*innen ohne einer medizinischen Ausbildung. Die OSCE Stationen konnten demnach die Leistungen medizinischen Personals tatsächlich als solche überprüfen und zwischen den Proband*innen mit besseren Fähigkeiten differenzieren.

Um die externe Validität zu steigern und zu prüfen, sollten jedoch weitere Studien in anderen Situationen und unter anderen Bedingungen durchführen werden^{120,171}.

5.6.7. Übertragbarkeit der Leistung der Proband*innen (Studierende) auf die Leistung von Ärzt*innen

Unter „Interaction of selection and treatment“ überprüfen Creswell et al., ob die Ergebnisse einer experimentellen Studie mit der Auswahl von bestimmten Proband*innen auf eine andere Personengruppe übertragen werden können^{120,171}. Im Falle der vorliegenden Studie ist dies relevant, da die Proband*innen ausschließlich Medizinstudierende und keine approbierten Ärzt*innen waren. Inwiefern das Verhalten und die Leistungen der Proband*innen auf das Verhalten von Ärzt*innen übertragbar sind, sollte daher kritisch hinterfragt werden.

Außerdem wurde die Studie zu zwei definierten Zeitpunkten im Studium umgesetzt. Im Sinne der „Interaction of history and treatment“ nach Creswell et al. sollte reflektiert werden, ob die Ergebnisse auf andere Zeitpunkte in der Zukunft, zum Beispiel wenn die Proband*innen das Studium abgeschlossen haben, übertragbar sind¹⁷¹.

Tatsächlich weisen Studien darauf hin, dass Erfahrungen im Umgang mit Unterbrechungen mögliche negative Effekt reduzieren können^{81,87,97} (s.a. Kapitel 5.2.5). Allerdings ist fraglich, ob im Studium entsprechende Erfahrungen gemacht werden oder eine Sensibilisierung für das Thema erfolgt. Der Umgang mit Unterbrechungen, Ablenkungen und Multitasking ist kein Bestandteil der Lehre im Medizinstudium. Da Medizinstudierende in ihren Praktika in der Regel kein eigenes Telefon zugewiesen bekommen, werden im Studium keine eigenen Erfahrungen bezüglich Unterbrechungen von ärztlichen Tätigkeiten durch Telefonanrufe gemacht. Die Erfahrungen können erst bei Berufsbeginn in den ersten Jahren der Berufstätigkeit gemacht werden. Daher kann davon ausgegangen werden, dass die Erkenntnisse aus der vorliegenden Studie auf unerfahrene Ärzt*innen in ihren ersten Berufsjahren übertragbar sind. Durch die große Anzahl von Proband*innen und das einschließen von zwei Semestern in die Studie sollten Unterschiede zwischen den einzelnen Proband*innen zudem ausgeglichen worden

sein. Weitere Studien mit Ärzt*innen als Proband*innen wären jedoch interessant, um die Ergebnisse zu überprüfen.

5.7 Fazit

In der vorliegenden Studie konnten negative Effekte von Unterbrechungen durch Telefonanrufe auf alltägliche ärztliche Tätigkeiten auch außerhalb des Operationssaals in einem experimentellen Studiendesign nachgewiesen werden.

Es zeigte sich eine schlechtere Leistung in der Bearbeitung der primären Aufgabe, das Vergessen sicherheitskritische Schritte, ein Effekt auf die Bearbeitungsdauer und auf die Gedächtnisleistung der Proband*innen.

Als relevante Faktoren für den Effekt von Unterbrechungen konnten Erfahrung im Umgang mit Unterbrechungen, der Zeitpunkt der Unterbrechung und die Komplexität der primären und sekundären Tätigkeit herausgearbeitet werden.

Damit konnte im Bereich Patient*innensicherheit eine Forschungslücke gefüllt sowie Erkenntnisse für die medizinische Praxis gesammelt werden. Anhand der Ergebnisse konnten zudem erste Empfehlungen für die Lehre erarbeitet werden, welche ergänzend zu einer systemischen Betrachtung berücksichtigt werden sollten.

Limitationen der Studie ergeben sich vor allem aus der Methodik. Das Risiko von Übertragungs-, Übungs- und Lerneffekt innerhalb der Studie ergeben sich aus dem time series, within-group Design. Außerdem ist zu berücksichtigen, dass die Erkenntnisse aus der OSCE-Prüfung möglicherweise nur eingeschränkt auf den klinischen Alltag übertragbar sind.

Angesichts der nachgewiesenen negativen Effekte erscheint eine Sensibilisierung für Effekte von Unterbrechungen und Ablenkungen und das Trainieren von nicht-technischen Fähigkeiten in der medizinischen Weiterbildung, unerlässlich.

Um das bisherige Wissen weiter auszubauen und Lösungsansätze gezielter ableiten zu können, sind jedoch weitere Studien nötig. Insbesondere eine weitere Evaluation möglicher Einflussfaktoren auf den Effekt von Unterbrechungen und Ablenkungen durch Telefonanrufe und eine Evaluation verschiedener Maßnahmen sollte in Zukunft erfolgen.

Nur so können Risiken für die Patientensicherheit in einem zunehmend von digitalen Hilfs- und Kommunikationsmitteln geprägten Gesundheitssystem identifiziert und reduziert werden.

6. Literaturverzeichnis

- 1 Toader JF, Kleinert R, Dratsch T, *et al.* Effect of phone call distraction on the performance of medical students in an OSCE. *BMC Med Educ* 2022; **22**. DOI:10.1186/S12909-022-03215-Y.
- 2 Gäckle A, Schumacher S, Thomas M. Leitfaden für eine geschlechtersensible Sprache 'überzeuGENDERE Sprache'. Die Gleichstellungsbeauftragte der Universität zu Köln V.i.S.d.P. Annelene Gäckle, 2021 https://gb.uni-koeln.de/e2106/e2113/e16894/20210709_Leitfaden_GGSprache_UzK_Webversion_ger.pdf (accessed Nov 19, 2025).
- 3 Universität zu Köln, Die Gleichstellungsbeauftragte. Geschlechtersensible Sprache. https://gb.uni-koeln.de/gendersensible_sprache/index_ger.html (accessed Nov 19, 2025).
- 4 Weigl M, Müller A, Zupanc A, Glaser J, Angerer P. Hospital doctors' workflow interruptions and activities: An observation study. *BMJ Qual Saf* 2011; **20**: 491–7.
- 5 Weigl M, Müller A, Vincent C, Angerer P, Sevdalis N. The association of workflow interruptions and hospital doctors' workload: A prospective observational study. *BMJ Qual Saf* 2012; **21**: 399–407.
- 6 Anderson CE, Nicksa GA, Stewart L. Distractions during resident handoffs: Incidence, sources, and influence on handoff quality and effectiveness. *JAMA Surg* 2015; **150**: 396–401.
- 7 Caird JK, Simmons SM, Wiley K, Johnston KA, Horrey WJ. Does Talking on a Cell Phone, With a Passenger, or Dialing Affect Driving Performance? An Updated Systematic Review and Meta-Analysis of Experimental Studies. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 2018; **60**: 101–33.
- 8 Caird JK, Willness CR, Steel P, Scialfa C. A meta-analysis of the effects of cell phones on driver performance. *Accid Anal Prev* 2008; **40**: 1282–93.
- 9 Weigl M, Müller A, Holland S, Wedel S, Woloshynowych M. Work conditions, mental workload and patient care quality: A multisource study in the emergency department. *BMJ Qual Saf* 2016; **25**: 499–508.
- 10 Weigl M, Müller A, Sevdalis N, Angerer P. Relationships of multitasking, physicians' strain, and performance: An observational study in ward physicians. *J Patient Saf* 2013; **9**: 18–23.
- 11 Augenstein T, Schneider A, Wehler M, Weigl M. Multitasking behaviors and provider outcomes in emergency department physicians: two consecutive, observational and multi-source studies. *Scand J Trauma Resusc Emerg Med* 2021; **29**. DOI:10.1186/s13049-020-00824-8.

- 12 Weigl M, Beck J, Wehler M, Schneider A. Workflow interruptions and stress atwork: A mixed-methods study among physicians and nurses of a multidisciplinary emergency department. *BMJ Open* 2017; **7**. DOI:10.1136/bmjopen-2017-019074.
- 13 Armendariz J, Tamayo C, Slade J, Belitskaya-Lévy I, Gray C, Allaudeen N. Interruptions to Attending Physician Rounds and Their Effect on Resident Education. *J Grad Med Educ* 2021; **13**: 266–75.
- 14 Wheelock A, Suliman A, Wharton R, et al. The impact of operating room distractions on stress, workload, and teamwork. *Ann Surg* 2015; **261**: 1079–84.
- 15 Weigl M, Antoniadis S, Chiapponi C, Bruns C, Sevdalis N. The impact of intra-operative interruptions on surgeons' perceived workload: an observational study in elective general and orthopedic surgery. *Surg Endosc* 2015; **29**: 145–53.
- 16 Weber J, Catchpole K, Becker AJ, Schlenker B, Weigl M. Effects of Flow Disruptions on Mental Workload and Surgical Performance in Robotic-Assisted Surgery. *World J Surg* 2018; **42**: 3599–607.
- 17 Antoniadis S, Passauer-Baierl S, Baschnegger H, Weigl M. Identification and interference of intraoperative distractions and interruptions in operating rooms. *Journal of Surgical Research* 2014; **188**: 21–9.
- 18 McMullan RD, Urwin R, Gates P, Sunderland N, Westbrook JI. Are operating room distractions, interruptions and disruptions associated with performance and patient safety? A systematic review and meta-analysis. *International Journal for Quality in Health Care* 2021; **33**. DOI:10.1093/intqhc/mzab068.
- 19 Westbrook JI, Raban MZ, Walter SR, Douglas H. Task errors by emergency physicians are associated with interruptions, multitasking, fatigue and working memory capacity: A prospective, direct observation study. *BMJ Qual Saf* 2018; **27**: 655–63.
- 20 Anzan M, Alwhaibi M, Almetwazi M, Alhawassi TM. Prescribing errors and associated factors in discharge prescriptions in the emergency department: A prospective cross-sectional study. *PLoS One* 2021; **16**. DOI:10.1371/journal.pone.0245321.
- 21 Wiegmann DA, ElBardissi AW, Dearani JA, Daly RC, Sundt TM. Disruptions in surgical flow and their relationship to surgical errors: An exploratory investigation. *Surgery* 2007; **142**: 658–65.
- 22 Sevdalis N, Undre S, McDermott J, Giddie J, Diner L, Smith G. Impact of Intraoperative Distractions on Patient Safety: A Prospective Descriptive Study Using Validated Instruments. *World J Surg* 2014; **38**: 751–8.
- 23 Weigl M, Müller A, Angerer P, Hoffmann F. Workflow interruptions and mental workload in hospital pediatricians: An observational study. *BMC Health Serv Res* 2014; **14**. DOI:10.1186/1472-6963-14-433.

- 24 Huhtala M, Geurts S, Mauno S, Feldt T. Intensified job demands in healthcare and their consequences for employee well-being and patient satisfaction: A multilevel approach. *J Adv Nurs* 2021; **77**: 3718–32.
- 25 Murji A, Luketic L, Sobel ML, Kulasegaram KM, Leyland N, Posner G. Evaluating the effect of distractions in the operating room on clinical decision-making and patient safety. *Surg Endosc* 2016; **30**: 4499–504.
- 26 Sujka JA, Safcsak K, Bhullar IS, Havron WS. Simulation-Based Testing of Pager Interruptions During Laparoscopic Cholecystectomy. *J Surg Educ* 2018; **75**: 1351–6.
- 27 Yang C, Heinze J, Helmert J, Weitz J, Reissfelder C, Mees ST. Impaired laparoscopic performance of novice surgeons due to phone call distraction: a single-centre, prospective study. *Surg Endosc* 2017; **31**: 5312–7.
- 28 Weigl M, Stefan P, Abhari K, *et al*. Intra-operative disruptions, surgeon's mental workload, and technical performance in a full-scale simulated procedure. *Surg Endosc* 2016; **30**: 559–66.
- 29 Walter SR, Li L, Dunsmuir WTM, Westbrook JI. Managing competing demands through task-switching and multitasking: a multi-setting observational study of 200 clinicians over 1000 hours. *BMJ Qual Saf* 2014; **23**: 231–41.
- 30 Westbrook JI, Ampt A, Kearney L, Rob MI. All in a day's work: An observational study to quantify how and with whom doctors on hospital wards spend their time. *Medical Journal of Australia* 2008; **188**: 506–9.
- 31 Weigl M, Müller A, Zupanc A, Angerer P. Participant observation of time allocation, direct patient contact and simultaneous activities in hospital physicians. *BMC Health Serv Res* 2009; **9**. DOI:10.1186/1472-6963-9-110.
- 32 Trbovich PL, Griffin MC, White RE, Bourrier V, Dhaliwal D, Easty AC. The effects of interruptions on oncologists' patient assessment and medication ordering practices. *J Healthc Eng* 2013; **4**: 127–44.
- 33 Pashler H, Jolicoeur P, Dell'Acqua R, *et al*. Task switching and multitask performance. In: Monsell S, Driver J, eds. *Attention and performance XVIII: Control of mental processes*. Cambridge, MA: MIT Press, 2000.
- 34 Douglas HE, Raban MZ, Walter SR, Westbrook JI. Improving our understanding of multi-tasking in healthcare: Drawing together the cognitive psychology and healthcare literature. *Appl Ergon* 2017; **59**: 45–55.
- 35 Salvucci DD, Taatgen NA, Borst JP. Toward a unified theory of the multitasking continuum: From concurrent performance to task switching, interruption, and resumption. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2009: 1819–28.

- 36 Healey AN, Sevdalis N, Vincent CA. Measuring intra-operative interference from distraction and interruption observed in the operating theatre. *Ergonomics* 2006; **49**: 589–604.
- 37 Healey AN, Primus CP, Koutantji M. Quantifying distraction and interruption in urological surgery. *BMJ Qual Saf* 2007; **16**: 135–9.
- 38 Keogh S, Laski D. A Concern for Intraoperative Distractions and Interference: An Observational Study Identifying, Measuring, and Quantifying Both within the Operating Theatre. *Surg Res Pract* 2021; **2021**. DOI:10.1155/2021/9910290.
- 39 Persoon MC, Broos HJHP, Witjes JA, Hendrikx AdJM, Scherpbier AJJM. The effect of distractions in the operating room during endourological procedures. *Surg Endosc* 2011; **25**: 437–43.
- 40 Allard J, Wyatt J, Bleakley A, Graham B. ‘Do you really need to ask me that now?’: A self-audit of interruptions to the “shop floor” practice of a UK consultant emergency physician. *Emergency Medicine Journal* 2012; **29**: 872–6.
- 41 Weigl M, Catchpole K, Wehler M, Schneider A. Workflow disruptions and provider situation awareness in acute care: An observational study with emergency department physicians and nurses. *Appl Ergon* 2020; **88**. DOI:10.1016/j.apergo.2020.103155.
- 42 Blocker RC, Heaton HA, Forsyth KL, et al. Physician, Interrupted: Workflow Interruptions and Patient Care in the Emergency Department. *Journal of Emergency Medicine* 2017; **53**: 798–804.
- 43 Ratwani RM, Fong A, Puthumana JS, Hettinger AZ. Emergency Physician Use of Cognitive Strategies to Manage Interruptions. *Ann Emerg Med* 2017; **70**: 683–7.
- 44 Coiera E, Tombs V. Communication behaviours in a hospital setting: An observational study. *Br Med J* 1998; **316**: 673–6.
- 45 Popovici I, Morita PP, Doran D, et al. Technological aspects of hospital communication challenges: an observational study. *International Journal for Quality in Health Care* 2015; **27 (3)**: 183–188.
- 46 Wu RC, Lo V, Morra D, et al. The intended and unintended consequences of communication systems on general internal medicine inpatient care delivery: A prospective observational case study of five teaching hospitals. *Journal of the American Medical Informatics Association* 2013; **20**: 766–77.
- 47 Tran K, Morra D, Lo V, Quan S, Wu R. The use of smartphones on General Internal Medicine wards: a mixed methods study. *Appl Clin Inform* 2014; **5**: 814–23.
- 48 Zarandona J, Cariñanos-Ayala S, Cristóbal-Domínguez E, Martín-Bezós J, Yoldi-Mitxelena A, Hoyos Cillero I. With a smartphone in one’s pocket: A descriptive cross-sectional study on smartphone use, distraction and restriction policies in nursing students. *Nurse Educ Today* 2019; **82**: 67–73.

- 49 Alameddine M, Soueidan H, Makki M, Tamim H, Hitti E. The Use of Smart Devices by Care Providers in Emergency Departments: Cross-Sectional Survey Design. *JMIR Mhealth Uhealth* 2019; **7**. DOI:10.2196/13614.
- 50 Cho S, Lee E. Distraction by smartphone use during clinical practice and opinions about smartphone restriction policies: A cross-sectional descriptive study of nursing students. *Nurse Educ Today* 2016; **40**: 128–33.
- 51 Nerminathan A, Harrison A, Phelps M, Alexander S, Scott KM. Doctors' use of mobile devices in the clinical setting: a mixed methods study. *Intern Med J* 2017; **47**: 291–8.
- 52 Terry DL, Terry CP. Smartphone Use and Professional Communication Among Medical Residents in Primary Care. *PRiMER* 2018; **2**. DOI:10.22454/primer.2018.766371.
- 53 Illiger K, Hupka M, von Jan U, Wichelhaus D, Albrecht U-V. Mobile technologies: expectancy, usage, and acceptance of clinical staff and patients at a university medical center. *JMIR Mhealth Uhealth* 2014; **2**. DOI:10.2196/mhealth.3799.
- 54 Hofer F, Haluza D. Are Austrian practitioners ready to use medical apps? Results of a validation study. *BMC Med Inform Decis Mak* 2019; **19**. DOI:10.1186/s12911-019-0811-2.
- 55 Franko OI, Tirrell TF. Smartphone app use among medical providers in ACGME training programs. *J Med Syst* 2012; **36**: 3135–9.
- 56 Boruff JT, Storie D. Mobile devices in medicine: A survey of how medical students, residents, and faculty use smartphones and other mobile devices to find information. *Journal of the Medical Library Association* 2014; **102**: 22–30.
- 57 Albrecht U-V, Afshar K, Illiger K, *et al*. Expectancy, usage and acceptance by general practitioners and patients: exploratory results from a study in the German outpatient sector. *Digit Health* 2017; **3**. DOI:10.1177/2055207617695135.
- 58 Yaman H, Yavuz E, Er A, *et al*. The use of mobile smart devices and medical apps in the family practice setting. *J Eval Clin Pract* 2016; **22**: 290–6.
- 59 Martin G, Janardhanan P, Withers T, Gupta S. Mobile revolution: A requiem for bleeps? *Postgrad Med J* 2016; **92**: 493–6.
- 60 Patel RK, Sayers AE, Patrick NL, Hughes K, Armitage J, Hunter IA. A UK perspective on smartphone use amongst doctors within the surgical profession. *Annals of Medicine and Surgery* 2015; **4**: 107–12.
- 61 El Hadidy TS, Alshafei AE, Mortell AE, Doherty EM. Smartphones in clinical practice: doctors' experience at two Dublin paediatric teaching hospitals. *Ir J Med Sci* 2018; **187**: 565–73.
- 62 Trafton JG, Altmann EM, Ratwani RM. A memory for goals model of sequence errors. *Cogn Syst Res* 2011; **12**: 134–43.

- 63 Couffe C, Michael GA. Failures due to interruptions or distractions: A review and a new framework. *American Journal of Psychology* 2017; **130**: 163–81.
- 64 Mark G, Gudith D, Klocke U. The cost of interrupted work: More speed and stress. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, 2008: 107–10.
- 65 Stelfox HT, Palmisani S, Scurlock C, Orav EJ, Bates DW. The 'To Err is Human' report and the patient safety literature. *Qual Saf Health Care* 2006; **15**: 174–8.
- 66 Institute of Medicine (US) Committee on Quality of Health Care in America, Kohn L, Corrigan J, Donaldson M. To Err is Human: Building a Safer Health System. Washington, DC: The National Academies Press (US), 2000 DOI:10.17226/9728.
- 67 Horrey WJ, Wickens CD. Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Hum Factors* 2006; **48**: 196–205.
- 68 Baldo N, Marini A, Miani M. Drivers' braking behavior affected by cognitive distractions: An experimental investigation with a virtual car simulator. *Behavioral Sciences* 2020; **10**. DOI:10.3390/BS10100150.
- 69 Palmiero M, Piccardi L, Boccia M, *et al.* Neural Correlates of Simulated Driving While Performing a Secondary Task: A Review. *Front Psychol* 2019; **10**. DOI:10.3389/fpsyg.2019.01045.
- 70 Sanbonmatsu DM, Strayer DL, Biondi F, Behrends AA, Moore SM. Cell-phone use diminishes self-awareness of impaired driving. *Psychon Bull Rev* 2016; **23**: 617–23.
- 71 McCurdie T, Sanderson P, Aitken LM. Traditions of research into interruptions in healthcare: A conceptual review. *Int J Nurs Stud*. 2017; **66**: 23–36.
- 72 Sanderson P, McCurdie T, Grundgeiger T. Interruptions in Health Care: Assessing Their Connection With Error and Patient Harm. *Hum Factors* 2019; **61**: 1025–36.
- 73 Grundgeiger T, Sanderson P. Interruptions in healthcare: Theoretical views. *Int J Med Inform* 2009; **78**: 293–307.
- 74 Rivera-Rodriguez AJ, Karsh B-T. Interruptions and distractions in healthcare: review and reappraisal. *Qual Saf Health Care* 2010; **19**: 304–12.
- 75 Gillespie BM, Chaboyer W, Fairweather N. Interruptions and Miscommunications in Surgery: An Observational Study. *AORN J* 2012; **95**: 576–90.
- 76 Yoong W, Khin A, Ramlal N, Loabile B, Forman S. Interruptions and distractions in the gynaecological operating theatre: irritating or dangerous? *Ergonomics*. 2015; **58**: 1314–9.
- 77 Allers JC, Hussein AA, Ahmad N, *et al.* Evaluation and Impact of Workflow Interruptions during Robot-assisted Surgery. *Urology* 2016; **92**: 33–7.
- 78 Catchpole K, Perkins C, Bresee C, *et al.* Safety, efficiency and learning curves in robotic surgery: a human factors analysis. *Surg Endosc* 2016; **30**: 3749–61.

- 79 Balint BJ, Steenburg SD, Lin H, Shen C, Steele JL, Gunderman RB. Do Telephone Call Interruptions Have an Impact on Radiology Resident Diagnostic Accuracy? *Acad Radiol* 2014; **21**: 1623–8.
- 80 Feuerbacher RL, Funk KH, Spight DH, Diggs BS, Hunter JG. Realistic distractions and interruptions that impair simulated surgical performance by novice surgeons. *Archives of Surgery* 2012; **147**: 1026–30.
- 81 Hsu KE, Man F-Y, Gizicki RA, Feldman LS, Fried GM. Experienced surgeons can do more than one thing at a time: effect of distraction on performance of a simple laparoscopic and cognitive task by experienced and novice surgeons. *Surg Endosc* 2008; **22**: 196–201.
- 82 Persoon MC, van Putten K, Muijtjens AMM, Witjes JA, Hendrikx AJM, Scherpbier AJJM. Effect of distraction on the performance of endourological tasks: a randomized controlled trial. *BJU Int* 2011; **107**: 1653–7.
- 83 Goodell KH, Cao CGL, Schwaitzberg SD. Effects of cognitive distraction on performance of laparoscopic surgical tasks. *Journal of Laparoendoscopic & Advanced Surgical Techniques* 2006; **16**: 94–8.
- 84 Moorthy K, Munz Y, Dosis A, Bann S, Darzi A. The effect of stress-inducing conditions on the performance of a laparoscopic task. *Surg Endosc* 2003; **17**: 1481–4.
- 85 Kilkenny JJ, Mrotz VJ, Khosa DK, Kerr CL, Denstedt JD, Singh A. The Impact of Distraction on Laparoscopic Skills in Veterinary Medical Students. *Veterinary Surgery* 2016; **45**. DOI:10.1111/vsu.12486.
- 86 Park J, Waqar S, Kersey T, Modi N, Ong C, Sleep T. Effect of distraction on simulated anterior segment surgical performance. *J Cataract Refract Surg* 2011; **37**: 1517–22.
- 87 Ghazanfar MA, Cook M, Tang B, Tait I, Alijani A. The effect of divided attention on novices and experts in laparoscopic task performance. *Surg Endosc* 2015; **29**: 614–9.
- 88 Evans CH, Schneider E, Shostrom V, Schenarts PJ. “Surgery interrupted”: The effect of multitasking on cognitive and technical tasks in medical students. *The American Journal of Surgery* 2017; **213**: 268–72.
- 89 Mellum ML, Vestergaard AH, Grauslund J, Vergmann AS. Virtual vitreoretinal surgery: effect of distracting factors on surgical performance in medical students. *Acta Ophthalmol* 2020; **98**: 378–83.
- 90 McGowan G, Jawaheer L, Young D, Yorston D. QUIET PLEASE! Effect of distraction on simulated posterior segment surgical performance. *Graefe’s Archive for Clinical and Experimental Ophthalmology* 2018; **256**: 519–23.
- 91 Cowan JB, Seeley MA, Irwin TA, Caird MS. Computer-simulated arthroscopic knee surgery: Effects of distraction on resident performance. *Orthopedics* 2016; **39**. DOI:10.3928/01477447-20160119-05.

- 92 Pluyter JR, Buzink SN, Rutkowski AF, Jakimowicz JJ. Do absorption and realistic distraction influence performance of component task surgical procedure? *Surg Endosc* 2010; **24**: 902–7.
- 93 Moorthy K, Munz Y, Undre S, Darzi A. Objective evaluation of the effect of noise on the performance of a complex laparoscopic task. *Surgery* 2004; **136**: 25–31.
- 94 Suh IH, Chien J-H, Mukherjee M, Park S-H, Oleynikov D, Siu K-C. The negative effect of distraction on performance of robot-assisted surgical skills in medical students and residents. *The International Journal of Medical Robotics and Computer Assisted Surgery* 2010; **6**: 377–81.
- 95 Cao S, Liu Y. Effects of concurrent tasks on diagnostic decision making: An experimental investigation. *IIE Trans Healthc Syst Eng* 2013; **3**: 254–62.
- 96 Monteiro SD, Sherbino JD, Ilgen JS, *et al.* Disrupting diagnostic reasoning: do interruptions, instructions, and experience affect the diagnostic accuracy and response time of residents and emergency physicians? *Acad Med* 2015; **90**: 511–7.
- 97 Li SYW, Magrabi F, Coiera E. A systematic review of the psychological literature on interruption and its patient safety implications. *Journal of the American Medical Informatics Association* 2012; **19**: 6–12.
- 98 Wu R, Rossos P, Quan S, *et al.* An evaluation of the use of smartphones to communicate between clinicians: A mixed-methods study. *J Med Internet Res* 2011; **13**. DOI:10.2196/jmir.1655.
- 99 Harden RM. Revisiting ‘Assessment of clinical competence using an objective structured clinical examination (OSCE)’. *Med Educ* 2016; **50**: 376–9.
- 100 Harden R, Downie WW, Stevenson M, Wilson GM. Assessment of Clinical Competence using Objective Structured Examination. *Br Med J* 1975; **1**: 447–51.
- 101 Barrows HS, Abrahamson S. The programmed patient: a technique for appraising student performance in clinical neurology. *J Med Educ* 1964; **39**: 802–5.
- 102 Barrows HS. An overview of the uses of standardized patients for teaching and evaluating clinical skills. AAMC. *Acad Med*. 1993; **68**: 443–53.
- 103 Vu NV, Barrows HS. Use of Standardized Patients in Clinical Assessments: Recent Developments and Measurement Findings. *Educational Researcher* 1994; **23**: 23–30.
- 104 Srisarajivakul N, Lucero C, Wang XJ, *et al.* Disruptive behavior in the workplace: Challenges for gastroenterology fellows. *World J Gastroenterol* 2017; **23**: 3315–21.
- 105 Khan KZ, Ramachandran S, Gaunt K, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part I: An historical and theoretical perspective. *Med Teach* 2013; **35**. DOI:10.3109/0142159X.2013.818634.

- 106 Bani-issa W, Al Tamimi M, Fakhry R, Tawil H Al. Experiences of nursing students and examiners with the Objective Structured Clinical Examination method in physical assessment education: A mixed methods study. *Nurse Educ Pract* 2019; **35**: 83–9.
- 107 Goldhamer MEJ, Cohen A, Brooks M, Macklin EA, Co JPT, Weinstein D. Use of an objective structured clinical exam (OSCE) for early identification of communication skills deficits in interns. *Med Teach* 2018; **40**: 40–4.
- 108 Laidlaw A, Salisbury H, Doherty EM, Wiskin C, UK Council for Clinical Communication in Undergraduate Medical Education. National survey of clinical communication assessment in medical education in the United Kingdom (UK). *BMC Medical Education* 2014; **14**. DOI:10.1186/1472-6920-14-10.
- 109 Sator M, Jünger J. Von der Insellösung zum Longitudinalen Kommunikationscurriculum - Entwicklung und Implementierung am Beispiel der Medizinischen Fakultät Heidelberg. *Psychother Psychosom Med Psychol* 2015; **65**: 191–8.
- 110 Simon SR, Bui A, Day S, Berti D, Volkan K. The relationship between second-year medical students' OSCE scores and USMLE Step 2 scores. *J Eval Clin Pract* 2007; **13**: 901–5.
- 111 Zyromski NJ, Staren ED, Merrick HW. Surgery residents' perception of the objective structured clinical examination (OSCE). *Curr Surg* 2003; **60**: 533–7.
- 112 Merrick HW, Nowacek G, Boyer J, Robertson J. Comparison of the objective structured clinical examination with the performance of third-year medical students in surgery. *The American Journal of Surgery* 2000; **179**: 286–8.
- 113 Sloan DA, Donnelly MB, Schwartz RW, Felts JL, Blue A V., Strodel WE. The Use of the Objective Structured Clinical Examination (OSCE) for Evaluation and Instruction in Graduate Medical Education. *Journal of Surgical Research* 1996; **63**: 225–30.
- 114 Farahat E, Javaherian-Dysinger H, Rice G, Schneider L, Daher N, Heine N. Exploring students' perceptions of the educational value of formative Objective Structured Clinical Examination (OSCE) in a nutrition program. *J Allied Health* 2016; **45**: 20–6.
- 115 Chaudhry SI, Holmboe E, Beasley BW. The state of evaluation in internal medicine residency. *J Gen Intern Med* 2008; **23**: 1010–5.
- 116 Martin IG, Jolly B. Predictive validity and estimated cut score of an objective structured clinical examination (OSCE) used as an assessment of clinical skills at the end of the first clinical year. *Med Educ* 2002; **36**: 418–25.
- 117 Müller S, Dahmen U, Settmacher U. Objective Structured Clinical Examination (OSCE) an Medizinischen Fakultäten in Deutschland – eine Bestandsaufnahme [Application of the Objective Structured Clinical Examination (OSCE) in German Medical Schools: An Inventory]. *Gesundheitswesen* 2018; **80**: 1099–103.

- 118 Wissenschaftsrat (2018). Neustrukturierung des Medizinstudiums und Änderung der Approbationsordnung für Ärzte | Empfehlungen der Expertenkommission zum Masterplan Medizinstudium 2020. Köln, 2018 <https://www.wissenschaftsrat.de/download/archiv/7271-18.html> (accessed Nov 21, 2025).
- 119 Hochberg MS, Kalet A, Zabar S, Kachur E, Gillespie C, Berman RS. Can professionalism be taught? Encouraging evidence. *Am J Surg* 2010; **199**: 86–93.
- 120 Creswell JW, Creswell DJ. Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research, 4th edn. Harlow, Essex: Pearson, 2014.
- 121 Woods B, Byrne A, Bodger O. The effect of multitasking on the communication skill and clinical skills of medical students. *BMC Med Educ* 2018; **18**. DOI:10.1186/s12909-018-1183-5.
- 122 Monk CA, Trafton JG, Boehm-Davis DA. The Effect of Interruption Duration and Demand on Resuming Suspended Goals. *J Exp Psychol Appl* 2008; **14**: 299–313.
- 123 Altmann EM, Trafton JG. Memory for goals: an activation-based model. *Cogn Sci* 2002; **26**: 39–83.
- 124 Trafton JG, Altmann EM, Brock DP, Mintz FE. Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human Computer Studies* 2003; **58**: 583–603.
- 125 Trafton JG, Altmann EM, Brock DP. Huh, what was I Doing? How People Use Environmental Cues after an Interruption. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 2005; **49**: 468–72.
- 126 Foroughi CK, Werner NE, McKendrick R, Cades DM, Boehm-Davis DA. Individual differences in working-memory capacity and task resumption following interruptions. *J Exp Psychol Learn Mem Cogn* 2016; **42**: 1480–8.
- 127 Paas F, Renkl A, Sweller J. Cognitive load theory and instructional design: Recent developments. *Educ Psychol* 2003; **38**: 1–4.
- 128 Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol Rev* 1956; **63**: 81–97.
- 129 Monsell S. Task switching. *Trends Cogn Sci* 2003; **7**: 134–40.
- 130 Rogers RD, Monsell S. Costs of a predictable switch between simple cognitive tasks. *J Exp Psychol Gen* 1995; **124**: 207–31.
- 131 Allport A, Wylie G. Task-switching: Positive and negative priming of task-set. In: Humphreys GW, Duncan J, Treisman A, eds. Attention, space, and action: Studies in cognitive neuroscience. Oxford University Press, 1999: 273–96.

- 132 Allport DA, Styles EA, Hsieh S. Shifting intentional set: Exploring the dynamic control of tasks. In: Umiltà C, Moscovitch M, eds. *Attention and performance 15: Conscious and nonconscious information processing*. Cambridge, MA, US: The MIT Press, 1994: 421–52.
- 133 Westbrook JI, Coiera E, Dunsmuir WTM, *et al.* The impact of interruptions on clinical task completion. *Qual Saf Health Care* 2010; **19**: 284–9.
- 134 Wickens CD. Multiple resources and mental workload. *Hum Factors*. 2008; **50**: 449–55.
- 135 Wickens CD. Multiple resources and performance prediction. *Theor Issues Ergon Sci* 2002; **3**: 159–77.
- 136 Engle RW, Tuholski SW, Laughlin JE, Conway ARA. Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *J Exp Psychol Gen* 1999; **128**: 309–31.
- 137 Yoder M, Schadewald D. The effect of a safe zone on nurse distractions, interruptions, and medication administration errors. *West J Nurs Res* 2012; **34**: 1068–9.
- 138 Hayes C, Jackson D, Davidson PM, Daly J, Power T. Calm to chaos: Engaging undergraduate nursing students with the complex nature of interruptions during medication administration. *J Clin Nurs* 2017; **26**: 4839–47.
- 139 Westbrook JI, Woods A, Rob MI, Dunsmuir WTM, Day RO. Association of interruptions with an increased risk and severity of medication administration errors. *Arch Intern Med* 2010; **170**: 683–90.
- 140 Bonafide CP, Miller JM, Localio AR, *et al.* Association between Mobile Telephone Interruptions and Medication Administration Errors in a Pediatric Intensive Care Unit. *JAMA Pediatr* 2020; **174**: 162–9.
- 141 Fang Q, Fang C, Li L, Song Y. Impact of sport training on adaptations in neural functioning and behavioral performance: A scoping review with meta-analysis on EEG research. *J Exerc Sci Fit* 2022; **20**: 206–15.
- 142 Voss MW, Kramer AF, Basak C, Prakash RS, Roberts B. Are expert athletes ‘expert’ in the cognitive laboratory? A meta-analytic review of cognition and sport expertise. *Appl Cogn Psychol* 2010; **24**: 812–26.
- 143 Chaddock L, Neider MB, Voss MW, Gaspar JG, Kramer AF. Do Athletes Excel at Everyday Tasks? *Med Sci Sports Exerc* 2011; **43**: 1920–6.
- 144 Anguera JA, Boccanfuso J, Rintoul JL, *et al.* Video game training enhances cognitive control in older adults. *Nature* 2013; **501**: 97–101.
- 145 Chiappe D, Conger M, Liao J, Caldwell JL, Vu KPL. Improving multi-tasking ability through action videogames. *Appl Ergon* 2013; **44**: 278–84.

- 146 Yu RWL, Chan AHS. Meta-analysis of the effects of game types and devices on older adults-video game interaction: Implications for video game training on cognition. *Appl Ergon* 2021; **96**. DOI:10.1016/J.APERGO.2021.103477.
- 147 Ball BH, Knight JB, Dewitt MR, Brewer GA. Individual differences in the delayed execution of prospective memories. *Q J Exp Psychol (Hove)* 2013; **66**: 2411–25.
- 148 Wolff J, Auber G, Schober T, *et al.* Work-time distribution of physicians at a German university hospital. *Dtsch Arztebl Int* 2017; **114**: 705–11.
- 149 Pezzolesi C, Ghaleb M, Kostrzewski A, Dhillon S. Is Mindful Reflective Practice the way forward to reduce medication errors? *International Journal of Pharmacy Practice* 2013; **21**: 413–6.
- 150 Matern LH, Farnan JM, Hirsch KW, Cappaert M, Byrne ES, Arora VM. A Standardized Handoff Simulation Promotes Recovery from Auditory Distractions in Resident Physicians. *Simul Healthc* 2018; **13**: 233–8.
- 151 Nguyen DL, Kamel P, Gomez EN. Preparing junior radiology residents for overnight call via peer-led, hands-on simulation. *Emerg Radiol* 2021; **28**: 589–99.
- 152 Spence H, Somasundram K, Biyani CS, Jain S. Training Nontechnical Skills in Ward Rounds to Improve Team Performance. *J Surg Educ* 2020; **77**: 921–30.
- 153 Somasundram K, Spence H, Colquhoun AJ, McIlhenny C, Biyani CS, Jain S. Simulation in urology to train non-technical skills in ward rounds. *BJU Int* 2018; **122**: 705–12.
- 154 Thomas I, Nicol L, Regan L, *et al.* Driven to distraction: a prospective controlled study of a simulated ward round experience to improve patient safety teaching for medical students. *BMJ Qual Saf* 2015; **24**: 154–61.
- 155 Lazarovici M, Trentzsch H, Prückner S. Human Factors in der Medizin. *Urologe* 2017; **56**: 97–113.
- 156 Catchpole K. Spreading human factors expertise in healthcare: Untangling the knots in people and systems. *BMJ Qual Saf*. 2013; **22**: 793–7.
- 157 Beasley JW, Wetterneck TB, Temte J, *et al.* Information chaos in primary care: Implications for physician performance and patient safety. *Journal of the American Board of Family Medicine* 2011; **24**: 745–51.
- 158 Leonard M, Graham S, Bonacum D. The human factor: the critical importance of effective teamwork and communication in providing safe care. *Qual Saf Health Care* 2004; **13**. DOI:10.1136/qhc.13.suppl_1.i85.
- 159 Behazin NS, Thompson M, Milad M, *et al.* Effects of the No Interruption Zone on Distraction Levels, Withdrawal Times, and Adenoma Detection Rates of Colonoscopy. *Dig Dis Sci* 2015; **60**: 966–70.
- 160 Hay JM, Barnette W, Shaw SE. Changing practice in gastrointestinal endoscopy reducing distractions for patient safety. *Gastroenterology Nursing* 2016; **39**: 181–5.

- 161 Patel R, Reilly K, Old A, Naden G, Child S. Appropriate use of pagers in a New Zealand tertiary hospital. *N Z Med J* 2006; **119**: U1912.
- 162 Fargen KM, O'Connor T, Raymond S, Sporrer JM, Friedman WA. An Observational Study of Hospital Paging Practices and Workflow Interruption Among On-Call Junior Neurological Surgery Residents. *J Grad Med Educ* 2012; **4**: 467–71.
- 163 Carlile N, Rhatigan JJ, Bates DW. Why do we still page each other? Examining the frequency, types and senders of pages in academic medical services. *BMJ Qual Saf* 2017; **26**: 24–9.
- 164 Ly T, Korb-Wells CS, Sumpton D, Russo RR, Barnsley L. Nature and Impact of Interruptions on Clinical Workflow of Medical Residents in the Inpatient Setting. *J Grad Med Educ* 2013; **5**: 232–7.
- 165 Levy JL, Freeman CW, Cho JK, Iyalomhe O, Scanlon MH. Evaluating the Impact of a Call Triage Assistant on Resident Efficiency, Errors, and Stress. *Journal of the American College of Radiology* 2020; **17**: 414–20.
- 166 Luu TB, Spiegelman LM, Nykin DM, *et al.* Implementation of an Electronic Health Record-Based Messaging System in the Emergency Department: Effects on Physician Workflow and Resident Burnout. *J Patient Saf* 2022; **18**. DOI:10.1097/PTS.0000000000000869.
- 167 Ernst AA, Weiss SJ, Reitsema JA. Does the Addition of Vocera Hands-Free Communication Device Improve Interruptions in an Academic Emergency Department? *South Med J* 2013; **106**: 189–95.
- 168 Williams LH, Drew T. Distraction in diagnostic radiology: How is search through volumetric medical images affected by interruptions? *Cogn Res Princ Implic* 2017; **2**. DOI:10.1186/s41235-017-0050-y.
- 169 Keren GB, Raaijmakers JGW. On between-subjects versus within-subjects comparisons in testing utility theory. *Organ Behav Hum Decis Process* 1988; **41**: 233–47.
- 170 Greenwald AG. Within-subjects designs: To use or not to use? *Psychol Bull* 1976; **83**: 314–20.
- 171 Creswell DJ, Cresswell JW. Research Design Qualitative, Quantitative, and Mixed Methods Approaches, 5th edn. Los Angeles: SAGE, 2018.
- 172 Hodges B. Validity and the OSCE. *Med Teach* 2003; **25**: 250–4.
- 173 Khan KZ, Gaunt K, Ramachandran S, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part II: Organisation & Administration. *Med Teach* 2013; **35**. DOI:10.3109/0142159X.2013.818635.
- 174 Hodges B, Regehr G, Hanson M, McNaughton N. Validation of an objective structured clinical examination in psychiatry. *Academic Medicine* 1998; **73**: 910–2.

- 175 Wilkinson TJ, Frampton CM, Thompson-Fawcett M, Egan T. Objectivity in objective structured clinical examinations: checklists are no substitute for examiner commitment. *Acad Med* 2003; **78**: 219–23.
- 176 Hodges B, McIlroy JH. Analytic global OSCE ratings are sensitive to level of training. *Med Educ* 2003; **37**: 1012–6.
- 177 Pierre RB, Wierenga A, Barton M, Branday JM, Christie CDC. Student evaluation of an OSCE in paediatrics at the University of the West Indies, Jamaica. *BMC Med Educ* 2004; **4**. DOI:10.1186/1472-6920-4-22.

7. Anhang

7.1 Abbildungsverzeichnis

Abbildung 1, Aufbau OSCE I.....	21
Abbildung 2, Aufbau OSCE II.....	22
Abbildung 3, Zeitlicher Ablauf Studienstation OSCE I.....	24
Abbildung 4, Zeitlicher Ablauf Studienstation OSCE II.....	24
Abbildung 5, Vergleich der erreichten Punktzahl OSCE I und OSCE II ¹	34
Abbildung 6, Vergleich der Erinnerungsleistung nach Namen, OSCE I.....	40
Abbildung 7, Vergleich der Erinnerungsleistung nach Namen, OSCE II.....	41
Abbildung 8, Punktezah der Proband*innen mit und ohne Ausbildung in OSCE I.....	43
Abbildung 9, Erreichte Punktezah nach Szenarien Kombination	47
Abbildung 10, Häufigkeitsverteilung der Szenarien Kombinationen	49

7.2 Tabellenverzeichnis

Tabelle 1, Checkliste OSCE I.....	28
Tabelle 2, Checkliste OSCE II.....	29
Tabelle 3, Materialien OSCE I.....	30
Tabelle 4, Materialien OSCE II.....	31
Tabelle 5, Unterschiede in der Bewertung der Checklisten Items zwischen der regulären Station und der Studienstation in OSCE I ¹	36
Tabelle 6, Unterschiede in der Bewertung der Checklisten Items zwischen der regulären Station und der Studienstation in OSCE II ¹	37
Tabelle 7, Anzahl an korrekt wiedergegebenen Informationen aus dem Telefonanruf in OSCE I und OSCE II ¹	39
Tabelle 8, Korrelation zwischen sportlicher Betätigung und erreichter Punktzahl	45
Tabelle 9, Korrelation zwischen der Geschwisteranzahl und erreichten Punktezah.....	45
Tabelle 10, Durchschnittlich erreichte Punktezah in den verschiedenen Szenarien.....	48

7.3 Materialien

7.3.1. Fragebogen

Fragebogen

Name, Vorname			
Teilnehmer-Nr.			
Geschlecht	männlich	weiblich	andere
Alter			
Studiensemester			
Muttersprache			
Nationalität			
Geschwister (ja/nein, Anzahl)			
Wie viele Famulaturen haben Sie bisher absolviert? (Anzahl in Wochen)			
Haben Sie vor dem Studium eine Ausbildung im medizinischen Bereich abgeschlossen, wenn ja welche?			
Machen Sie Sport? Falls ja wie oft in der Woche und wie lange?			
Spielen Sie Videospiele?			
Falls ja: Wann haben Sie angefangen? (Jahr)			
Falls ja: Wieviel Stunden am Tag/in der Woche? - Als Kind/Während der Schulzeit - Während des Studiums - Aktuell			
Haben Sie eine eigene Spielkonsole zu Hause?			

7.3.2. Materialien OSCE I

7.3.2.1. Fallvignette OSCE I reguläre Station

OSCE- Station: Venenverweilkanüle

Setting:

Herr Klütsch soll eine Infusion erhalten. Sie werden auf die Station gerufen, um Ihm einen peripher venösen Zugang zu legen und die Infusion anzuhängen.

Aufgabenstellung:

Legen Sie eine Venenverweilkanüle und hängen sie anschließend die Infusion an.

Achten Sie dabei auf Ihr hygienisches und strukturiertes Vorgehen.

Insgesamt haben sie 5 Minuten Zeit!

Viel Erfolg!

7.3.2.2. Fallvignette OSCE I Studienstation

Studien-Station: Venenverweilkanüle

Setting:

Herr Klütsch soll eine Infusion erhalten. Sie werden auf die Station gerufen, um ihm einen peripher venösen Zugang zu legen und die Infusion anzuhängen.

Aufgabenstellung:

Legen Sie eine Venenverweilkanüle und hängen sie anschließend die Infusion an.

Achten Sie dabei auf Ihr hygienisches und strukturiertes Vorgehen.

Bitte beachten Sie: Sie sind diensthabender Arzt der Station der Viszeralchirurgie. Sie bekommen ihr Diensttelefon angereicht und können hierüber jederzeit angerufen werden. Falls Sie angerufen werden nehmen Sie den Anruf bitte an.

Insgesamt haben Sie 5 Minuten Zeit.

Bitte gehen Sie nach dem Lesen der Vignette direkt in den Raum.

Viel Erfolg!

Die Teilnahme ist freiwillig. Das Ergebnis dieser Station geht nicht in das Ergebnis der Prüfung ein. Wir danken für die Teilnahme.

7.3.2.3. Aufteilung Blutzuckerwerte und Patientennamen

Liste der Namen und BZ-Werte für die Telefonanrufe

	Dienstag	Mittwoch	Donnerstag
Gruppe 1, 10:10-11:05	Frau Fischer, BZ 220	Herr Mutlu, BZ 205	Herr Weber, BZ 210
Gruppe 2, Start 11:05- 12:00	Herr Weber, BZ 210	Frau Fischer, BZ 220	Herr Mutlu, BZ 205
Gruppe 3, Start 12:00- 12:55	Herr Mutlu, BZ 205	Herr Weber, BZ 210	Frau Fischer, BZ 220
<i>Mittagspause</i>			
Gruppe 4, Start 13:25- 14:20	Herr Weber, BZ 210	Frau Fischer, BZ 220	Herr Weber, BZ 210
Gruppe 5, Start 14:20- 15:15	Herr Mutlu, BZ 205	Herr Weber, BZ 210	Herr Mutlu, BZ 205
Gruppe 6, Start 15:15-16:10	Frau Fischer, BZ 220	Herr Mutlu, BZ 205	Herr Weber, BZ 210
<i>Nachmittagspause</i>			
Gruppe 7 Start 16:30-17:25	Herr Weber, BZ 210		Frau Fischer, BZ 220
Gruppe 8 Start 17:25-18:20	Herr Mutlu, BZ 205		Herr Mutlu, BZ 205

7.3.3. Materialien OSCE II

7.3.3.1. Fallvignette OSCE II reguläre Station

OSCE- Station: Kurzanamnese und Untersuchung des Abdomens

Setting:

Patient / Patientin mit Bauchbeschwerden wird vom Hausarzt
eingewiesen,
T 38.4° C, Leukozyten 14.000.

Aufgabenstellung:

- Erheben Sie kurz eine gezielte Anamnese in Bezug auf die geschilderte abdominelle Problematik.
- Untersuchen Sie das Abdomen und beschreiben Sie die einzelnen Schritte Ihrer Untersuchung. Erläutern Sie dabei die zu erwartenden Befunde.
- Geben Sie eine abschließende Bewertung der erhobenen Befunde.

Es befindet sich ein zweiter Simulationspatient im Raum. Die beiden wechseln sich von Prüfling zu Prüfling ab. Bitte ignorieren Sie den zweiten Patienten.

Insgesamt haben Sie 5 Minuten Zeit.

Viel Erfolg!

7.3.3.2. Fallvignette OSCE II Studienstation

Studien-Station: Kurzanamnese und Untersuchung des Abdomens

Setting:

Patient / Patientin mit Bauchbeschwerden wird vom Hausarzt
eingewiesen,
T 38.4° C, Leukozyten 14.000.

Aufgabenstellung:

- Erheben Sie kurz eine gezielte Anamnese in Bezug auf die geschilderte abdominelle Problematik.
- Untersuchen Sie das Abdomen und beschreiben Sie die einzelnen Schritte Ihrer Untersuchung. Erläutern Sie dabei die zu erwartenden Befunde.
- Geben Sie eine abschließende Bewertung der erhobenen Befunde.
- **Sie sind diensthabender Arzt der Station der Viszeralchirurgie. Sie bekommen ihr Diensttelefon angereicht und können hierüber jederzeit angerufen werden. Falls Sie angerufen werden nehmen Sie den Anruf bitte an.**

Der Prüfer befindet sich im Raum, bitte ignorieren Sie ihn.
Insgesamt haben Sie 5 Minuten Zeit.

Bitte gehen Sie nach dem Lesen der Vignette direkt in den Raum.

Viel Erfolg!

Die Teilnahme ist freiwillig. Das Ergebnis dieser Station geht nicht in das Ergebnis der Prüfung ein. Wir danken für die Teilnahme.

7.3.3.3. Aufteilung Blutzuckerwerte und Patientennamen

Liste der Namen und BZ-Werte für die Telefonanrufe

Gruppe und Startzeit (kann sich Nachmittags ggf. verschieben)	Dienstag (9.7)	Mittwoch (10.7)	Donnerstag (11.7)
Gruppe 1 - 8:30	Herr Weber, BZ 210	Herr Mutlu, BZ 205	Frau Fischer, BZ 220
Gruppe 2 - 10:10	Frau Fischer, BZ 220	Herr Weber, BZ 210	Herr Mutlu, BZ 205
Gruppe 3 - 11:50	Herr Mutlu, BZ 205	Frau Fischer, BZ 220	Herr Weber, BZ 210
<i>Mittagspause</i>	<i>Mittagspause</i>	<i>Mittagspause</i>	<i>Mittagspause</i>
Gruppe 4 - 13:50	Frau Fischer, BZ 220	Herr Mutlu, BZ 205	
Gruppe 5 – 15:30	Herr Weber, BZ 210	Frau Fischer, BZ 220 (Die Gruppe entfällt ggf.)	

7.3.3.4. Szenarien Kombination

Liste Szenarien (reguläre Station – Studienstation)

Gruppe und Startzeit (kann sich Nachmittags ggf. verschieben)	Dienstag (9.7)		Mittwoch (10.7)		Donnerstag (11.7)	
	Reguläre Station	Studienstation	Reguläre Station	Studienstation	Reguläre Station	Studienstation
Gruppe 1 - 8:30	Appendizitis	Divertikulitis	Cholezystitis	Appendizitis	Appendizitis	Cholezystitis
Gruppe 2 - 10:10	Divertikulitis	Appendizitis	Divertikulitis	Cholezystitis	Divertikulitis	Appendizitis
Gruppe 3 - 11:50	Cholezystitis	Appendizitis	Cholezystitis	Appendizitis	Cholezystitis	Appendizitis
Mittagspause	Mittagspause		Mittagspause		Mittagspause	
Gruppe 4 - 13:50	Appendizitis	Cholezystitis	Cholezystitis	Divertikulitis		
Gruppe 5 – 15:30	Divertikulitis	Cholezystitis				

8. Vorabveröffentlichungen von Ergebnissen

Toader, Justus & Kleinert, Robert & Dratsch, Thomas & Fettweis, Louisa & Jakovljevic, Nadja & Graupner, Martina & Zeeh, Moritz & Kroll, Anna & Fuchs, Hans & Wahba, Roger & Plum, Patrick & Bruns, Christiane & Datta, Rr. (2022). Effect of phone call distraction on the performance of medical students in an OSCE. BMC Medical Education. 22. 10.1186/s12909-022-03215-y.